A5030-4U

THE "LAKE WOBEGON" REPORT

How Public Educators Cheat on Standardized Achievement Tests

John Jacob Cannell

FRIENDS FOR EDUCATION

Copyright © 1989 by Friends of Education. All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without written permission from the author.

Printed in the United States of America by Cottonwood Printing, Albuquerque, NM

Cover art and illustrations by Carolyn Kinsman

About Friends for Education
Friends for Education is a non-profit, citizen's group working for improvements in America's public schools. Contributions to Friends for Education are tax-deductible and sorely needed.

Additional copies of this report can be obtained for $15.00 from Friends for Education, 600 Girard Blvd. NE, Albuquerque, NM 87106. We have no staff, so all orders must be prepaid. Purchase orders cannot be accepted.

I dedicate this book to my wife,
Ina,
whose love, encouragement, and patience
is remarkable.

ACKNOWLEDGEMENTS

I wish to express my appreciation and respect for The Kettering Family Foundation of Dayton, Ohio, which awarded a $25,000 grant that made this study possible. They supported our efforts to bring accountability to public education with full realization of the controversy surrounding our work. Eighty other prominent American foundations denied our request for support.

I also wish to express my admiration for the faculty of the Department of Psychiatry of the University of New Mexico, especially Dr. Eberhart Uhlenhuth, Dr. Robert Hendron, and Dr. Robert Kellner. They have been most supportive in their attitude and constructive in the comments, in spite of the controversial nature of this publication.

However, the reader should not infer that this study necessarily reflects the opinions of either the Department of Psychiatry of the University of New Mexico, or the staff and trustees of The Kettering Family Foundation. The opinions expressed in this study are solely those of the author.

In addition, I wish to express my appreciation to Julie Dunham, Amy Nixon, and Jamie Sercie for their help in obtaining data and revising the manuscript. Special thanks also to Liesl Meyers of The Desktop Press and to Ginger Griffin of Le Griffe Editorial, both located in Albuquerque, New Mexico. Their help was professional, expert, and invaluable.

Finally, I have borrowed for my title Garrison Keillor's magical place name "Lake Wobegon." In his inspired Swiftian way, like a modern Gulliver, Keillor has gently and humorously exposed the weakness, foibles, and failures of the current American scene.

REVIEWERS

Before publication, this study was reviewed by testing and measurement authorities, adult and child psychiatrists, and concerned citizens. Their contributions improved the report immensely. Some of the reviewers took alternative viewpoints of both the problem and its solutions. Their names and addresses are listed below. The reader should not infer that the following persons agree with our conclusions or endorse our solutions.

Mr. Roy Casto
Public School Teacher
Woodrow Wilson High School
Briarwood Road
Beckley, WV 25801

Mr. Walter Faithorn
Retired Business Executive
Stuart-Warner Corporation
3800 Underwood Street
Chevy Chase, MD 20815

Mr. Chester Finn, Jr.
Director
Educ. Excellence Network
1112 Sixteenth Street, NW
Suite 500
Washington, DC 20036

Dr. Thomas Fisher
Director, Testing & Evaluation
Department of Education
506 Knott Bldg.
Tallahassee, FL 32399

Professor Walter Haney
College of Education
Boston College
500 McGuinn Hall
Chesnut Hill, MA 02167

Dr. Allan Hartman
Director of Research
Massachusetts Department of Education
1385 Hancock Street
Quincy, MA 02169

Dr. Robert Hendron
Director
Division of Adolescent and Child Psychiatry
University of New Mexico School of Medicine
2400 Tucker, NE
Albuquerque, NM 87131

Professor Robert Linn
College of Education
University of Colorado
Box 249
Boulder, CO 80309

Professor William Mehrens
Education Department
Michigan State University
462 Erickson Hall
East Lansing, MI 48824-1034

Dr. John Mendenhall
Department of Psychiatry
University of New Mexico School of Medicine
2400 Tucker, NE
Albuquerque, NM 87131

Professor Jason Millman
Education Department
Cornell University
301 Roberts Hall
Ithica, NY 14853-5901

Professor E. Uhlenhuth
Residency Director
Department of Psychiatry
University of New Mexico School of Medicine
2400 Tucker, NE
Albuquerque, NM 87131

CONTENTS

	Acknowledgements	ii
	Reviewers	iii
	Preface	v
	Introduction	1
Chapter 1	**How Can Everyone Be Above Average**	6
Chapter 2	**"High Stakes" Tests**	9
Chapter 3	**Cheating?**	12
Chapter 4	**Test Security in the 50 States**	19
Chapter 5	**Deceptive Testing Practices**	24
Chapter 6	**Misleading Reporting Methods**	27
Chapter 7	**Suspicious Patterns in Test Scores**	30
Chapter 8	**Inexpensive Steps to Avoid Cheating**	33
Chapter 9	**The Effect of Inflated Test Scores**	39
Appendix I	Nationally Normed Test Results of The 50 States	42
Appendix II	Survey of Test Security Practices in The 50 States	100
Appendix III	Other Indicators of State Achievement	103
	Selected References	110
	About The Author	114

PREFACE

"Education is a companion which no misfortune can depress, no crime can destroy, no enemy can alienate, no despotism can enslave. At home a friend, abroad an introduction, in solitude a solace, and in society an ornament. It chastens vice, it guides virtue, it gives, at once, grace and government to genius. Without [illegible] what is man? A splendid slave, a reasoning savage."

Joseph Addison: The Spectator, November 6, 1711

I am neither a professional educator or a testing expert. I am a physician. For ten years I had a general practice in West Virginia. I am now a resident psychiatrist at the University of New Mexico. I have seen and continue to see many adolescent patients with serious self-esteem problems. Drug abuse, delinquency, teenage pregnancy, teenage depression and, most sadly, teenage suicide are the common presenting signs and symptoms of low self-esteem.

As a General Practitioner in West Virginia, I referred these patients to a clinical psychologist for an evaluation which routinely included grade level testing for academic ability. The results were continually unsettling; many adolescents with self-esteem problems were sitting in seventh grade general studies classrooms with third grade reading abilities.

When I asked school officials about these adolescents' school records, I was told these patients scored well on the school's "standardized" tests, including reading. The results of tests administered in the schools were contrary to the independent testing administered outside the school system, testing which placed them years behind their class.

A Case in Point

Kim (not her real name) was a 15 year old ninth grader who came to me complaining of morning nausea, breast tenderness and a three month cessation of her period. Her mother related she was well adjusted until the end of the third grade when she began

having nightmares, difficulty sleeping, moodiness, school phobia, and periods of frank depression. Symptoms had been present intermittently for the last five years, but disappeared during the summers.

The mother stated her daughter experimented with drugs in the fifth grade, had a history of truancy starting in the sixth grade, began having sexual intercourse in the seventh grade, and was put on probation for petty larceny and possession of drugs in the eighth grade. She had seen a child psychologist for five months at age twelve, and again for three months at age fourteen. Both parents were employed full time and there was no history of mental illness, alcoholism, or child abuse in the family.

A physical exam revealed the patient had an enlarged uterus and a serum pregnancy test confirmed she was three months pregnant. Kim seemed relieved when informed of my findings. She planned on having and keeping the baby. The father of the child was "going steady" with another girl.

I referred her to an obstetrician and a child psychologist. The psychologist diagnosed a clinical depression and psychotherapy was initiated. During therapy Kim related that school had become a source of acute embarrassment for her beginning in the third grade. She had nightmares about being asked to read in class, and suffered anxiety attacks when asked to do math problems at the blackboard. Her lack of ability was the object of derisive jokes by her classmates. She had internalized her lack of basic skills as a lack of self-worth. Therapy was directed at reinforcing her self-esteem.

Although, her IQ tested at 112, her reading ability tested at a fourth grade level and her math ability at a third grade level. Three months later, Kim terminated psychotherapy. She quit school at age 16, and delivered a healthy baby boy three months later. Through her parents I learned she was unmarried, living at home, and working at a local fast food restaurant to support the baby.

I had requested and obtained her official school record. It indicated that Kim had never been offered remediation, never been required to attend summer school, and had never been retained. Her fourth grade teacher had considered and then rejected retention because the teacher didn't want Kim's "self-esteem to be injured." Her school had administered a standardized achievement test, the Comprehensive Test of Basic Skills (CTBS), in the third and sixth grade. It indicated Kim's "total basic skills" were slightly "above the national norm."

Naturally, I began to wonder about the contradiction that these radically different results present. How could so many children test below average on independent testing but do well on their official school achievement test? The U.S. Department of Education could not give me any information. They do not oversee "standardization" or verify the norming process for the commercial achievement tests used in America's public schools. In fact, the U.S. Department of Education was not able to tell me any of the state scores.

"Above The National Average"

Just about this time, West Virginia and Kentucky announced their state total battery scores were "above the national average" on the same McGraw Hill CTBS achievement test. These two states have some of the highest illiteracy and poverty rates in the nation and some of the lowest college entrance and Armed Services Vocational Aptitude Battery scores. It did not make sense that they were above the national average on the McGraw Hill CTBS achievement test.

I became increasingly troubled. I decided to present myself to a test publisher as a superintendent of schools from a small southern Virginia school district. I called a publisher and expressed interest in purchasing this company's standardized achievement test. I explained that our district was considering changing tests and stated that the local school board was very interested in improving test scores.

Almost immediately, I was talking to a saleswoman who implied that our district's scores would be "above average" if we bought one of their "older" tests! She further intimated that our scores would go up every year, as long as we "didn't change tests."

What was an "older" test? How could she know that our district would be above the national average? The district whose name I used is a poor rural southern Virginia district. How could she guarantee yearly improvements as long as we "didn't change tests." She couldn't know if this district's schools were improving or not.

I had been aware of rumors about cheating in the schools. Many teachers privately told me that school personnel changed students' answer sheets after the test, gave students more than the allotted time, used the exact test questions to review for the test, or made copies of the test to give to their students. Many teachers complained that administrators forced them to teach items known to be on the test, claiming they could not get a promotion without producing high test scores.

Friends For Education

In 1984, I formed the Friends for Education, a tax exempt, non-profit educational watchdog group, and I became an education activist, working for improvements in public schools through "accountability." We obtained additional test results for local schools such as college entrance scores and military test results. They painted a very different picture of local school quality than did the commercial, "nationally normed," norm-referenced tests used locally.

We organized meetings about school quality, but when we criticized local school problems we were immediately confronted with official "standardized achievement test" scores, scores that were always above average. The state claimed that the tests were only used as "instructional aids," but the state's own press releases indicated that the scores actually served as an internal framework for

measuring school quality. The schools used them to measure student achievement, and, through the student's scores, they also measured teacher competency, school quality, and administrators' effectiveness. **Virtually all the local schools were testing "above the national average," even in the poorest districts of West Virginia!**

We realized that central office school administrators routinely studied these achievement tests to "align the curriculum" with the test questions. Even without overt cheating, it was no wonder that scores went up every year and that everybody was testing above average. The schools used the same exact questions year after year-questions which were "aligned into the curriculum," and which teachers then taught, often unwittingly, often with full awareness.

Forewarned Is Forearmed

It became clear why the saleswoman could guarantee scores would go up every year as long as we didn't change tests. The schools and the publishers they had under contract were jointly claiming that scores were improving because schools were improving. However, the actual process under way was increasingly efficient revelation to students, before their test, of the questions that would be on their test. The schools, in cooperation with their contract publishers, were teaching the students the answers in advance.

The schools then compared their current scores to the scores of a norm group tested in the past by the commercial publishers. Unlike the currently tested students, the norm group took the test "cold"; that is, the norm group didn't have the advantage of having their curriculum "aligned with the test." With norm-referenced tests, everybody could test "above the national norm."

No legitimate standardized testing service allows school personnel to know test content in advance. Publishers and local school authorities claimed the scores were improving because the schools were improving, even though evidence indicated "teaching the test" was responsible for improved scores. In addition, legitimate standardized tests only allow 50 percent of the students to test "above average."

I now understood why so many of my low self-esteem patients testing above the national norm without knowing how to read. I understood why every school I looked at was testing above the national norm while report after national report deplored the condition of American education.

"Lake Wobegon Effect"

I decided to survey all 50 states to see if any states were testing below the publisher's "national norm." Friends for Education had not yet obtained any outside funding so I, my nurse, lab technician, and X-ray technician called and wrote letters to state

education departments requesting test information. After reviewing the responses, we could not find one state below average at the elementary level on their total battery of scores! After obtaining results from more than 3500 school districts, we concluded that 70% of American school children, 90% of American school districts, and all 50 states were testing above the publishers "national norm" on commercial norm-referenced achievement tests.

We subsequently published *How All 50 States Are Above the National Average*. Our study showed that some of the poorest, most desperate school districts in the nation are able to pacify the press, parents, and school board by testing "above the national norm" on one of these commercial "Lake Wobegon" achievement tests. East Saint Louis, New York City, Boston, Akron, Columbus, Toledo, Trenton, East Orange and Paterson, New Jersey—even Harlan County Kentucky, and McDowell County, West Virginia—were all testing "above the national norm" at the elementary level on one of these commercial achievement tests, in spite of widespread illiteracy in their classrooms.

After I published the "Lake Wobegon" study, I started receiving letters detailing the extent of unethical testing practices in our schools. Some teachers openly admitted cheating. Others were concerned that if they didn't cheat, they would look bad compared to the teachers who did. All the teachers complained that cheating is encouraged by school administrators.

I am convinced that the current American epidemic of teenage pregnancy, depression, drug use, delinquency, and teenage suicide is partially related to the low standards and the low expectations so evident in America's public schools. School officials blame these problems on single parent families, parental apathy, and permissive child-rearing. Undoubtedly, many of these present day realities do detrimentally affect children, but so do present day school policies.

Our organization is also convinced that schools with high expectations and high standards could help the self-esteem of many of these children instead of further damaging it. However, school boards and state legislators cannot improve American schools until they have accurate information on local achievement. For that reason, we hope to be able to continue updating our "Lake Wobegon" report, with the belief that better public school accountability will eventually mean better public schools.

John Jacob Cannell M.D.
Albuquerque, New Mexico
August, 1989

This report was made possible by a generous grant from The Kettering Family Foundation of Dayton, Ohio. However, we have exhausted our funds and future reports will depend on additional funding.

INTRODUCTION

"A little learning is a dangerous thing;
Drink deep, or taste not the Pierian spring:
There shallow draughts intoxicate the brain,
And drinking largely sobers us again."

Alexander Pope: An Essay on Criticism, 1711

Americans are frustrated with the never-ending "school crisis." Numerous special committees have issued report after report recommending solutions. America has quadrupled spending on education over the last 16 years [$934 per student per year in 1972, $3,977 per student per year in 1988] (1,2). "School choice" is the latest solution, just as increased spending, merit pay, school-based management, and teacher competency testing were solutions a few years ago. But these don't seem to solve the problem.

Americans have grown accustomed to reports which decry the state of American education. We elect "education presidents." Television specials regularly alarm us about widespread young adult illiteracy. Volunteers tutor American adults on the basic skills in all 50 states. American business pays for a 36-page supplement in *The Wall Street Journal* which condemns American schools as "obsolete," calling for "a complete overhaul, not just more tinkering" (3).

The U.S. Secretary of Education agrees. In a recent speech at a conference of The Education Commission of the States, Secretary Cavazos stated: "Our schools, overall, have not changed their pattern of operation. They show little or no progress in measures of educational achievement." After 15 years of school reform, why haven't American schools improved?

Two Different Messages

Could it be that Americans hear two different messages about school quality: a pessimistic national message and an optimistic local message? Americans believe that public schools are a disgrace nationally, but they also believe that their local schools are doing a good job.

The national message is consistently pessimistic. It comes from complaints by American business (3), surveys of young American adult illiteracy (4), the National Assessment of Educational Progress (5), international comparisons of student achievement (6), and from nationwide results on certain standardized tests such as the Scholastic Aptitude Test (SAT), College Board, and the American College Test (ACT).

For example, the National Assessment of Educational Progress and similar studies find:

- Just five percent of seventeen year old American high school students can read well enough to understand and use information found in technical materials, literary essays, and historical documents.
- Only 37 percent can find France on a map and just 25 percent can find Massachusetts.
- Only six percent of seventeen year old American high school students can use basic algebra to answer the following question: "Christine borrowed $850 for one year from the bank. If she paid 12 percent simple interest on the loan, what is the total amount she repaid?"
- Only twenty-five percent of seventeen year old American high school students know when Lincoln was president, and just twenty percent know what "Reconstruction" was about.

As the "school crisis" approaches its 15th birthday, national achievement tests such as the National Assessment of Educational Progress continue to send their pessimistic messages. For example, the recent 36-page *Wall Street Journal* supplement on American education stated: "The tide of educational mediocrity, to paraphrase the 1983 report, *A Nation at Risk*, hasn't appreciably turned. Scholastic Aptitude Test scores fell again last year. Only marginal gains have been made in mathematics and science test scores" (3). This pessimistic national message is based on tests that are administered to children all over the country, under standardized conditions and under strict security.

However, Americans, especially school board members and state legislators, hear a second and very different message: one that is uniformly optimistic. For example, the same *Wall Street Journal* article went on to praise one local school district, Rochester, N.Y., by referring to their California Achievement test, stating: "scores are going up" (3). In fact, Rochester's "standardized test" scores have gone up steadily, and their elementary scores are "above the national norm." These are norm-referenced achievement tests scores, tests that are administered locally under conditions the commercial publishers prescribe, and which compare local achievement to a "national norm" supplied by the publishers.

Norm-Referenced Yardsticks: True or False?

American school board members almost invariably depend on one of six different norm-referenced tests to compare local achievement with national achievement: the California Achievement Test, the Stanford Achievement Test, the Metropolitan Achievement Test, the Science Research Associates Test, the Comprehensive Test of Basic Skills, and the Iowa Test of Basic Skills. In the last 15 years, these six tests have become the principal local yardsticks, the local internal report cards of American education.

The tests that give us the pessimistic national message, the National Assessment of Educational Progress, the College Board, and the ACT are not used to assess local school achievement. Even college entrance scores and Armed Services Vocational Aptitude tests, which are available as local district reports, are seldom obtained by local school boards. Instead, American school boards and state legislators depend on "standardized" norm-referenced achievement tests to measure local achievement. Local officials have unwittingly assumed that "a standardized test is a standardized test."

However, norm-referenced tests were designed years ago to be instructional aids, not "accountability" yardsticks. They told teachers which subjects, and students, needed the most help. Now, parent report forms, "classroom performance profiles," school achievement reports, district educational brochures, and state "accountability" publications rank individual, class, school, district, and state scores as above or below the "national norm" of the 50th percentile.

Norm-referenced tests are also used to detect improvements or declines in achievement over time and thus have been used by school boards, state legislators, and the press to assess local school improvements as well as overall school quality, teacher and administrator competency, and program effectiveness. Moreover, state "accountability" publications are testimony to the fact that public school administrators themselves now rely on norm-referenced test scores to measure school quality, not to aid instruction.

For example, a June 1, 1989, Alabama press release states: "We exceeded the national norms in grades one and two, and we are continuing to show improvement in the "above average" category at all other grade levels" (7). The 1989 *South Carolina Statewide Testing Program Summary Report* depicts graphs with South Carolina achievement towering above the "national median (8). It summarizes: "The scores on the CTBS/U continue to make visible the educational improvements which have taken place in South Carolina over the past seven years." Most states have similar publications; virtually all states claim their achievement is above the "national norm."

Feel Good Tests

The vast majority of American school districts have shown steady, often dramatic, improvement in norm-referenced test scores over the last 15 years. Even, Chicago, labeled "the worst school system in the nation" by the former U.S. Secretary of Education, has shown steady gains on their Iowa Test of Basic Skills. Chicago scores will soon exceed the "national norm" (9). School boards across the country have watched with satisfaction as local "standardized" commercial test scores have soared to record levels.

Local "standardized" test scores convince parents, school board members, and state legislators that local schools are doing a good job. For example, American elementary students recently lagged well behind other countries in international comparisons of reading and math achievement. The only category where Americans surpassed other countries was in their opinion of local American schools. Ninety-one percent of the mothers of these students rated their child's school "good or excellent," in spite of the fact their children lagged far behind their international colleagues in basic skills!

In fact, scores on local norm-referenced achievement tests have improved so dramatically that by early 1988 all 50 states were testing above the publisher's national norm (10). Our first report, *How All 50 States Are Above the National Average*, documented how 70 percent of American elementary children, 90 percent of American school districts, and all 50 states were testing above the publisher's "national norm" on norm-referenced achievement tests, instead of the expected 50 percent.

"Lake Wobegon" Tests

The Associated Press labeled our first report, "The Lake Wobegon Report," after Garrison Keillor's mythical town in Minnesota where "all the women are strong, all the men are good looking, and all the children are above average" (11). "Lake Wobegon" norm-referenced, elementary achievement tests have led almost all local school districts in the nation to believe they are achieving above the national average because they test above the publisher's national norm on a "standardized" test. More importantly, these tests have convinced local officials, like in Rochester and Chicago, that local schools are improving.

But, why, as recently stated by the U.S. Department of Education, "have other national and international assessment programs not reported the kind of high achievement" found on norm-referenced achievement tests (12)? **Because, there are critical differences between the norm-referenced achievement tests used locally and the standardized tests used nationally, such as the College Board, the ACT, and NAEP.** This report will underline these differences, differences that give parents, press, educators, school board members, and state legislators falsely optimistic and dangerously comforting beliefs about local school achievement.

Suspect Testing Practices

Our latest survey of the 50 states documents that the vast majority of American school systems continue to score above the publisher's national average, at both elementary and secondary grades, instead of the expected 50 percent. "Lake Wobegon" achievement scores are being reported as "above the national average" by some of the worst school systems in the nation.

This report also includes the first state-by-state survey of test security practices ever published. **We found that high scores on "Lake Wobegon" tests are often caused by lax test security, nonstandard testing practices, deceptive statistics, and misleading impressions, not improved achievement. Most upsetting, the report concludes that outright cheating by American educators on "Lake Wobegon" tests of school achievement is common.**

This report will give responsible officials a method to evaluate local achievement scores to see if testing irregularities or cheating should be suspected. **Most importantly, the report will suggest practical and inexpensive ways that school board members and state legislators can correct the problems.**

CHAPTER 1

How Can Everyone Be Above Average?

"What is it men cannot be made to believe!"

Thomas Jefferson: Letter to Richard H. Lee, 1786

The U.S. Department of Education called our first study, *How All 50 States Are Above the National Average*, a "major public service" (12) and commissioned researchers at the University of Colorado and UCLA to repeat our first study. Within the last few months, these investigators confirmed Friends for Education's basic findings.

First Study Scrutinized

At the 1989 meeting of The American Educational Research Association, Dan Koretz, the senior social scientist at the Rand Corporation, stated that our first study was found to be "clearly right"(13). Researchers at the University of Colorado studied "Lake Wobegon" achievement tests and concluded that "the overall percent of students above the national median is greater than 50 in all the elementary grades in both reading and math for each of the three years studied" (14). They further concluded that "achievement gains reported to the public based on standardized achievement tests appear to be exaggerated" (15).

However, the researchers found some technical flaws in our first study, such as not including secondary scores and not being able to obtain consistent score reporting methods from the states. Linn found that approximately 60 percent of the elementary children in the nation score above the national median in reading, and 65 percent in math, instead of the expected 50 percent.

Linn's figures were lower than Friends for Education's finding that 70 percent of American children were testing above the national average on norm-referenced tests, but as Linn points out, the differences between his findings and

ours "is largely attributable to the differences in definitions" (14). We used composite or total battery scores above 50, and Linn used reading scores above 50 to determine the percentage of children testing above the publisher's national norm. For reasons we will explain later, reading is always the lowest score in districts that cheat.

Most Recent Findings

Friends for Education has once again obtained the latest official standardized test scores from all 50 states. After correcting the technical problems, we still find that the vast majority of local school districts in the nation are testing above the national average, instead of the expected 50 percent. "Above average" scores are being reported by some of the poorest school systems in our nation.

The most recent scores, together with technical details of our latest survey, can be found in Appendix I. **This year, we find that 48 of the 50 states are scoring "above the national norm." In addition, we conclude that more than 90 percent of the 15,000 elementary school districts, and 80 percent of the secondary school districts in the nation are scoring "above the national norm" on "standardized achievement tests," instead of the expected 50 percent.**

Legitimate "Standardized Tests"

Standardized tests the public are familiar with, like the College Board, the National Assessment of Educational Progress, and the ACT, compare currently tested students with all other currently tested students by computing a yearly national average. In addition, most use a type of scaled score that allows year-to-year comparisons with past acnievement as well. With such scores, gains or losses in achievement over time can be measured, but only 50 percent of students are allowed to be above the current national average.

However, the "Lake Wobegon" norm-referenced achievement tests compare current student achievement to the achievement of a norm group tested in the past. Current national averages are not computed. All students who score higher than the publisher's old norm group are "above the national average." Thus, "Lake Wobegon" tests make it statistically possible for 100 percent of students, and school districts, to be "above the national norm."

Norm groups are selected and tested by the commercial publishers without government supervision or regulation. These norm groups are supposed to represent an average group of students tested under conditions similar to conditions used in current testing programs. However, unlike currently tested students, norm groups are "tested cold" without any prepping on the exact test questions and without having their curriculum "aligned" with the test questions (16).

Overstated Achievement

The public assumes that above the "national norm" means above the national average. They assume that their child, their school, their district, or their state is achieving at a level that exceeds current average national achievement. This is not an unreasonable assumption because the word "norm" is equated with the word "average" in the common lexicon—as evidenced in every dictionary we have consulted (e.g., *Webster's New World*).

State legislators and school board members assumed that "standardized" tests are all standardized in the same manner. This confusion has resulted in a false impression of local school quality. **After all, how can local schools be part of a "Nation at Risk" if they are above the national average and improving on their standardized tests?**

Dan Koretz, analyzing our first study in the 1988 summer edition of *American Educator*, concludes: "In my opinion, there can be no doubt that current norm-referenced tests overstate achievement levels in many schools, often by a large margin" (16). That "overstatement" occurs primarily at the elementary level and has convinced many Americans that secondary school, not elementary school, is where our nation is at risk.

CHAPTER 2

"High Stakes" Tests

> I hate by-roads in education. Education is as well known, and has long been as well known, as ever it will be."
>
> *Samuel Johnson: Bosewell's Life, 1775*

Professor Jim Popham at UCLA coined the term "high stakes" for tests that have consequences. When teachers feel judged by the results, when parents receive reports of their child's test scores, when tests are used to promote students, when test scores are widely reported in newspapers, then the tests are "high stakes."

Tests that were once used only as instructional aids now assess class achievement, school achievement, and district achievement through student's scores. As such, the class results then become the teacher's score, the school's results become the principal's score, and the district's results become the superintendent's score.

Individual student's results are sent home to parents via computerized report forms that describes student achievement in terms of the national norm of 50. Class and school scores are reported to teachers and principals in a similar manner. Scores are reported to the school board and the state legislature by means of an "accountability" report.

Media Attention

Educators are under tremendous pressure to raise test scores, usually by their state legislature. Teachers' skills will be questioned unless their students score high. In some states merit pay is based on "Lake Wobegon" test scores, and districts can be proclaimed "educationally bankrupt" and placed in state receivership for low test scores (15).

Teachers are compared to each other based on their students' test scores, and principals are judged on their schools' score. Superintendents are also under tremendous pressure to produce high scores. School boards see good scores as good schools and improving scores as improving schools. High test scores are the proof that many superintendents need to get a raise or just to keep their jobs.

Test scores are also released to the press, usually by a press conference. As can be seen in the state's press releases and "accountability" publications, "Lake Wobegon" tests are now used by educators to measure school quality through aggregated achievement scores. School administrators know "Lake Wobegon" scores are published in the local paper; scores are often the subject of four or five news articles a year.

Tests Measure Educators, Not Students

In most states, "Lake Wobegon" tests are much "higher stakes" for educators than for students. In fact, low student scores have almost no consequences at all for individual students. Only a few states, like New Jersey and Nevada, use norm-referenced scores to identify children needing remediation.

Many states claim that their testing program is a "low stakes" test, that is, the test is only used to help children or for instructional purposes. For example, in a letter to me, the West Virginia superintendent of schools claimed that: "Our standardized testing program is designed primarily for instructional improvement rather than comparisons."

However, West Virginia's Department of Education press releases, informational brochures, and "accountability" reports tell a different story. As their most recent state achievement report makes clear, the testing was mandated by the West Virginia legislature "to measure achievement and progress within public and non-public West Virginia schools."

Testing Pressure Cooker

In most states, "Lake Wobegon" scores are published in the newspapers on a school-by-school and district-by-district basis. In some states, districts can be declared "educationally bankrupt" for consistently low test scores. I have seen principal's test scores pinned on their wall.

Georgia, New Jersey, California and Texas are other examples of states with "high stakes" testing programs. In Georgia, "Lake Wobegon" test scores can affect how far a teacher will advance up a "career ladder" program. In New Jersey, elementary schools must report the number of children that perform below a state-mandated score on one of the "Lake Wobegon" tests. In California, school-by-school test scores are widely printed in the media and affect real estate prices. For a time, Texas based merit pay on "Lake Wobegon" scores. Forty of the 50 states have similar "high stakes" testing (15).

Dr. Walter Haney, the respected professor of testing and measurement at Boston College, concluded: "When testing programs have large consequences for kids, teachers, or schools, considerable pressure can build up to boost test scores regardless of the method. Scandals over cheating on tests hit the front pages with remarkable regularity" (17).

CHAPTER 3

Cheating?

"O, that deceit should dwell
In such a gorgeous palace."

Shakespeare: Romeo and Juliet, III, 1596

Just because the motivation to cheat is present, one cannot assume that cheating is taking place. We must still decide if the dramatic increases on "Lake Wobegon" tests are real or fake. Do the improved scores reflect improved school achievement or irregular testing practices?

In response to an advertisement we placed in *Education Week* this spring, Friends for Education received letters from teachers around the country detailing the extent of cheating in our public schools. As one might expect it is concentrated in those states with "high stakes" tests.

Teachers Blow the Horn on Cheating

Consider this letter from a teacher in Tennessee:

"Dear Dr. Cannell:

As a teacher, I have been repeatedly astounded in recent years concerning what is going on in testing. I think you would be absolutely flabbergasted if you knew how much cheating now takes place on the various achievement and basic skills tests in public schools.

One of our elementary schools was recently named by Instructor *as one of its top ten elementary schools in the nation. Yet it is common knowledge among the teachers, principals, and supervisors that this school took twice as long as usual to administer the Stanford Achievement Test because*

they spent the morning teaching the test and the afternoon giving it. That school went from near the bottom to the top in test results in our county in one year.

It appears that there are teachers in every school now who cheat. The most common way is to teach the test. There are pressures in states like Tennessee, where advancement is competitive, that are almost irresistible.

It is not that difficult to find out if cheating is taking place on yearly achievement tests. Many teachers and administrators take cheating very lightly now. I am from the older generation and will retire in a few years. I consider this a national scandal of enormous proportions that should come out.

Sincerely,"

Is this a "scandal of enormous proportions"? The letter is hearsay, the teacher offers no proof. Consider another letter:

"Dear Dr. Cannell:

I am a teacher in the state of Arkansas and am appalled at what is happening in the name of education here. Our governor has decided to revolutionize education in the state of Arkansas and has devised a test that students must pass or be retained in the 8th grade.

It doesn't matter if the educators are good teachers. They are only rewarded for good test scores. For example, teachers can get copies of any of these tests because they are the ones who administer the tests. How well an individual student, class, or school does on the test depends more on the dishonesty of the teacher, rather than what the student knows. Students pass the test unfairly because of teachers who cheat.

Sincerely,"

Again, no proof. Is it just the bitterness of a lazy teacher or something else? Another letter:

"Dear Dr. Cannell,

I am a kindergarten teacher and I have watched other kindergarten teachers' IOWA scores show most of their students score in the ninetieth percentile. I naively thought it would not take the administration long to figure out that they were teaching the test.

At one point I questioned our principal about the high

scores. I was told, "confidentially", that not only did he know that some teachers were teaching the test, he also knew that some teachers walked around during the test, pointing out to their students which questions needed to be "rechecked."

Although I knew my scores were making me look bad, I held out for four years. Finally, during my fifth year of teaching kindergarten, I, too, stooped to teaching the tests. I hardly spent more than one week teaching the test during the 1985-86 school year. My scores jumped to the 90th percentile, too. The administration said nothing. The 1986-87 school year I went back to not teaching the test. The scores, once again, fell tremendously.

I have read your report concerning nationally normed elementary achievement testing with great interest. For years I have watched what was going on in our public schools. My questions and concerns have usually been met with no interest. My lower IOWA test scores have led some parents to question my teaching skills. In the end, I have concluded, that there is really nothing I can do about this fraud. I would appreciate your comments on the matter.

Sincerely,"

This teacher admits cheating, but the only proof she can offer is that she herself cheated. We received many similar letters, some were unsigned, but most made it clear that they wanted to remain anonymous.

Note: These letters have been edited for clarity, space requirements, and confidentiality.

Investigating Cheating

What hard evidence about cheating on standardized achievement tests exists? We could find only a few school districts in the country that ever undertook a random audit of schools for cheating and then reported their findings.

In the course of investigating charges of cheating by 23 schools on the Iowa Test of Basic Skills, Chicago's testing officials selected 17 schools that were above suspicion to serve as controls for the 23 schools that were suspected of cheating (18). Chicago's school officials confirmed cheating in 70 percent of the 23 suspect schools. Officials also discovered cheating in 12 percent of the 17 control schools, and they suspected cheating in additional control schools! The study was very conservative, and investigators admitted they "may have underestimated the extent of cheating at some schools."

Officials in Chicago detected cheating by having outside proctors admini-

ster an equivalent but different form of the Iowa Test of Basic Skills. The suspect principals and teachers never handled the audit test booklet or the answer sheets. This type of audit should catch most types of cheating. However, if the same test is used for audit purposes that was initially administered by the school, then teaching the test, which is the easiest way to cheat, would be entirely missed.

Besides teaching the test, the authors concluded that educators cheated in Chicago by allowing students to exceed the publisher's time limits, providing the students with the correct answers during the test, and altering the answer sheets after the test. No disciplinary action was taken against any of the suspect principals or teachers at any of the 19 schools where cheating was confirmed. Even school personnel at the four schools that submitted altered answer sheets escaped disciplinary action.

From Suspicious Erasures to Allowing More Time

California is one of only a few states that systematically audits testing to detect altered answer sheets. Cheating is detected by optically scanning for a suspiciously high number of erasures on answer booklets for the California Assessment Program. Booklets with more than two standard deviations above the normal number of erasures are then checked by hand to see if the answers have all been changed from wrong to right instead of the expected distribution of changes.

In the last three years, 50 elementary schools in the state of California have been caught cheating by officials from the California State Department of Education (19). This year the electronic erasure scanning system found suspicious erasures in 18 different elementary schools in Los Angeles alone.

Remember, this system only detects one type of cheating: educators in a back room erasing students' wrong answers and then marking the correct answers. It can not detect easier kinds of cheating such as allowing students more than the allotted time, helping students with the answers during the test, or teaching the test questions.

Other schools in California have been caught cheating in different ways (19). Bandini Elementary School in San Pedro was caught when officials noticed that their reading scores jumped from the 5th percentile to the 94th percentile in one year. In this case, test security was breached by school personnel who were suspected of obtaining the test and then teaching test items directly to students in advance of the test.

Backing Into Discoveries of Cheating

In Austin, Texas, cheating has been reported on a number of occasions (20). However, the authors of that study admitted they have never run an audit to find out the actual incidence of cheating and further admitted: "We back into most of our discoveries of cheating." Some of the cases the authors "backed into" include:

1] Students reported that their teacher had given out copies of the test for homework the week before the test. Discipline consisted of transferring the teacher to another school.

2] Administrators noticed that many former students of one particular teacher lost ground the next year on their standardized tests. Cheating was suspected, but no action was taken with this Teacher of the Year honoree, even after outside test proctors confirmed that the teacher's test scores had been artificially inflated.

3] The district's lowest achieving school became the highest achieving school within a single year. Teachers were found to be reading out the answers to students during testing. No disciplinary action was taken.

4] A teacher photocopied the vocabulary words on the test and passed them out to students before the test. A letter of censure was placed in the teacher's file.

5] A teacher gave a copy of the test to her daughter so she could use it to prepare for the test. There was no report of any disciplinary action taken.

6] It was reported that teachers told special education students and other low achievers to stay home on the day of testing. No disciplinary action was taken.

The director of testing for the Austin Independent School District was forced to conclude: "**Teachers cheat when they administer standardized tests to students.** Not all teachers, not even very many of them but enough to make cheating a major concern for all of us who use test data for decision making" (20).

"Conspiracy of Silence"

The testing coordinator of a school district in Oregon reported finding copies of the district's standardized test "posted in some rooms" (21). He also reported teachers were using the test to prep their classes. The district was so concerned about security that it decided to have the test coordinator start administering the test instead of teachers.

When the Dallas Board of Education adopted a merit pay plan that awarded teachers bonuses for high classroom test scores, the superintendent of Dallas Public Schools admitted he needed to do something to allay "public perception of gross cheating throughout the school system" (22). The president of the Classroom Teachers of Dallas admitted that Dallas teachers were teaching the test, but he stated: "We were just following orders from administrators." The president of the Dallas Federation of Teachers agreed, stating: "Everyone in every school knows someone who is cheating. Teachers get pressure from the principal" (23).

There have been reports of cheating on the Regents Exam in New York for years (24). Most recently, the *New York Post* decided that cheating was so wide-

spread it printed the answer key on the front page (25). The *Post* said it showed how easy it was to illegally obtain the Regent's answer key, a key that is supposed to be secure in the principals' offices. The *Post's* editor stated: "I was astounded that the state Education Department...was going to proceed with a fiction."

In Montgomery County, Maryland, 291 test scores were disqualified after testing irregularities were found, but no disciplinary action was taken (26). Virginia has also reported cheating. Two administrators and three teachers were reprimanded for cheating after their high test scores were found to be caused by teaching the test (27).

The former Virginia state testing director concluded that cheating occurred "a lot more than the department was willing to admit. I really have some bad feelings about some improper procedures that took place. Jack Davis (the state school superintendent) would not let us do anything. **It was almost like a conspiracy of silence"** (27).

Administrators Benefit

One computerized statistical technique used to detect cheating is called cluster analysis. When applied at a school in Dallas it detected isolated instances of probable irregularities, mostly at elementary schools (28). Teachers were suspected of obtaining test items in advance of the test and coaching students on the correct answers.

Researchers at Virginia Polytechnic Institute concluded that cluster analysis, "is straightforward and effective and has been available for a number of years, but it has had very little use" (28). The authors concluded that, "School systems avoid quality checks on their testing simply because they fear revelations of deficiencies."

Much of the unethical behavior has become so commonplace in our public schools that educators no longer see it as cheating. Typical of such behavior is a Los Angeles teacher who walked around the classroom during the examination looking for students who had marked wrong answers on their California Assessment Program test answer booklets (19). He admitted pointing out their incorrect responses and instructing students to "fix their mistakes." "I don't see it as wrong," he said.

One Los Angeles school official was asked what he thought about increasing test scores by having teachers teach their class the exact vocabulary words on the test. "I don't see that as a significant problem. That would be evidence of growth" (29).

In Washington, D.C., the Associate Superintendent for Instruction and the administrator in charge of the Washington school system's "test preparation program" admitted encouraging teachers to teach students the answers (30). The administrator admitted that the district's educators "even begin to tell them (students) some of the things on the test." The authors concluded that "this practice is

controversial and raises ethical concerns."

A better idea of the administrators' role in cheating comes from a businessman in Baltimore, a Harvard graduate who taught in an inner-city school for nine years (31). He became so disgusted with "test score falsification" that he quit teaching and went into business. He writes:

"If the Iowa Test of Basic Skills contained a reading comprehension selection concerning the Roman toga, teachers were instructed to make sure that the children were familiar with all the concepts and vocabulary words that they might encounter in the passage on the Roman toga. Perhaps most amusing, the principal would come on the intercom each morning to discuss one or two vocabulary items from the ITBS vocabulary subtest: "Good morning, boys and girls, this morning we are going to have a little `chat.' Do you know what a `chat' is? That's right, it's a little talk." Chat, of course, was one of the items on the ITBS vocabulary subtest."

One of the reasons that administrators do not take disciplinary measures against teachers found to be cheating on standardized tests is that administrators are the ones who benefit most from rising scores. In many cases administrators have convinced teachers that cheating is not only acceptable but desirable.

CHAPTER 4

Test Security In The 50 States

"The great secret of education is
to direct vanity to proper objects."

Adam Smith: The Theory of Moral Sentiments, 1759

Just as the public assumes that "above the national norm" means "above the national average," they also assume standardized achievement tests are administered under standardized conditions—conditions that prevent cheating. The public assumes that norm-referenced tests measure achievement in such a way that improving scores means improving overall achievement. Legitimate "high stakes" standardized tests take great care to see that such an association can be made.

Unfair Advantages

If a student steals a math test and studies the questions in advance of the test, the test will not be a random sampling of his math ability. He only has to study the test to do well; his classmates have to study the entire math book to do as well.

Likewise, if a teacher obtains a math test and then teaches his class the exact questions, then the test is not a random sampling of his class's achievement. His class only has to study the test to do well; other classes have to study the entire math book to do as well.

If the school continues to use the same test, class scores will rise as the teacher drills more of his class on the correct answers. Overall basic skills may decrease as the teacher spends more time teaching the exact test questions and less time teaching basic skills (16). When teachers spend valuable class time drilling students with the exact test questions, it is always at the expense of the regular curriculum.

In the same manner, if a superintendent of schools obtains a normed-referenced test in advance, and then narrows his district's curriculum to focus on test

content, he is assured of inflated scores compared to the norm group who studied a general curriculum.

Legitimate "High Stakes" Test Security

Legitimate "high stakes" standardized tests never allow educators to see the actual test in advance. Scores will be inflated if questions are known and somehow emphasized in the curriculum. Legitimate tests like the ACT, the National Assessment of Educational Progress, and the College Board seal the test booklets or the packages in which they are delivered; all use outside test proctors, employed by the testing service, to administer the test. No one but the student can look at the questions; in fact, students are often told not to take the test if the seal is broken.

More importantly, legitimate "high stakes" standardized tests change most questions every year by substituting different but equivalent questions. For example, the College Board, the National Assessment of Educational Progress, and the ACT rotate enough questions every year so next year's test will not be compromised, even if this year's test is stolen. Next year's test will assess the same content and skills by using different questions. This is the most important method that legitimate standardized tests use to ensure year-to-year test security.

"Lake Wobegon" Test Security

Our survey of state test security (Appendix II) found that "Lake Wobegon" tests are administered without similar security measures. Although most states had written test security and test procedure policies, most of these policies were totally inadequate; many policies were only two or three sentences that did not address the cheating problem. Rhode Island's test security policies, as outlined in their *1988 Testing Coordinator's Handbook*, are typical:

1. Store materials in rooms or cabinets that are locked, and that are not readily accessible to large numbers of other people.
2. Check all materials as you receive them to verify counts; have counts verified again when material are returned for storage.
3. Keep all extra test materials in the secure location when they are not in use.

Security Policy Problems

The problem with such security policies are numerous. First, the principals are often responsible for securing the "locked rooms or cabinets." As recently seen in California, principals have strong incentives to cheat, especially if their school may be declared "educationally bankrupt," or embarrassed in the local newspaper with low test scores (19).

Second, making sure that tests "are not readily accessible to large numbers of other people," is simply silly in the age of photocopiers and fax machines. Just one person can breech security. As New York recently discovered on its Regents exam, even shipping tests in locked metal chests will not ensure security (25).

Third, policies which focus on accounting for exact numbers of test booklets or forbid teachers to copy booklets without limiting the time and opportunity educators have access to booklets, are ineffectual.

As recently pointed out by Dr. Lori Shepard, "Lake Wobegon" tests are exquisitely sensitive to being taught (15). An educator that remembers only one test question, and subsequently teaches it to his class, can raise his classes' score from the 48th percentile to the 55th percentile!

Basic Security Lacking

Our survey of state test security policies found that most policies do not address these problems. Most districts deliver open boxes of unsealed "Lake Wobegon" tests to their schools weeks before the test is to be given. Teachers are often given unsealed test booklets days before they are scheduled to administer the test. Only sixteen states claimed that they require test administrators to receive the test booklets no earlier than the day of administration, but four others said they planned to institute this precaution.

We found that only seven states forbid teachers to read the test booklets except as needed during the actual administration. However, six states plan to institute this basic security procedure. We found five states that actually encourage teachers to take the test well in advance of administering it.

We found only 12 states that seal their test booklets, either individually (four states), or by shrink wrapping them in class size packets (10 states). Three states claimed they were planning to seal their test booklets. Most states allow the unsealed tests to stay in the schools for prolonged periods of time.

Tests could easily be administered in the early fall, at the beginning of school, when the results would reflect last year's achievement, not this year's. Fall testing reduces the pressure on teachers to teach the test because they see low fall scores as a reflection on last year's teachers, not themselves. We found only eight states that take this simple step to reduce the incentive to cheat, although three more states are planning a fall administration.

Testing Not Monitored

Most states have no method of ensuring that tests are administered under standardized conditions. Test proctors are used in only seven states. Only 17 states have empowered their testing directors to send out monitoring teams to randomly monitor test administration procedures, and two states are planning to, but few state testing directors have the resources, staff, or political power to monitor testing effectively.

Teachers could easily be forbidden from administering these "high stakes" tests to their own class. Instead, they could be required to switch with other teachers on the day of testing and administer the test to another class at another grade level. We could not find any states that used this simple method of reducing the opportunity for teachers to cheat.

Although most publishers offer equivalent versions of the same test, and states could rotate equivalent forms of the test, we found only 18 states that did so, and three that are planning to rotate questions. Test rotation could be achieved at no additional expense by purchasing, for example, one half of the required number of tests as form A and the other half as form B, and then rotate the forms. Such simple measures are seldom taken. In most states the teachers administer the same test to their own classes, with the same unchanging questions, year after year.

Administrators Have Opportunity to Cheat

Superintendents are responsible for test security in their own district and thus have access to "Lake Wobegon" tests. In many cases the tests are stored in the district office all year, where they could easily be copied.

In most states, the principals collect and store the answer sheets until the central office picks them up; this storage time provides a prime opportunity for answer sheets to be altered, as was done in 31 Los Angeles elementary schools last year (19). Only six states use technological devices to detect such alterations, although the technology is widely available and inexpensive. Four additional states plan to start scanning answer sheets for alterations.

Answer sheets can be analyzed for suspicious groupings of responses by different cluster analysis computer programs. This procedure can detect classes who may have been prepped on the test, by flagging classrooms where all the students marked the same response on the same question. When this happens repeatedly, it may indicate the teacher is prepping his class on the test. We could find only five states that routinely analyze answer sheets for such groupings.

Good Security in Nation's Largest District

Many districts claim they are too large to institute these test security measures. However, New York City, the largest school district in the country, is one of the few school districts that have most of these security policies in place (32).

For example, New York City testing officials deliver shrink wrapped class size packets of test booklets to schools shortly before testing, give booklets to teachers on the day of testing, and forbid teachers to read test booklets except as needed during administration.

They also send out unannounced test monitoring teams to assure packets have not been broken early, rotate one-third of the questions every year on their

customized Metropolitan Achievement Test, and analyze some answer sheets for cluster variance.

Blaming Teachers Instead of Administrators

Instead of adequate test security policies, one state, South Carolina, has made it a felony for teachers to knowingly teach test items. However, South Carolina delivers unsealed boxes of CTBS tests to schools days or even weeks before testing is scheduled to begin (see Appendix II). They do not monitor their CTBS assessment, and they allow teachers to read the test booklets in advance of administering it. Teachers know the test, which has been used for six years in a row, will be used again next year.

Laws that make it a felony to teach the test are somewhat fatuous and do very little. They do allow school administrators to claim they take cheating very seriously, but the absence of meaningful test security policies seem designed to make it easy for teachers to cheat. Again, it is the administrators, as well as teachers, who benefit when test scores go up.

"Lake Wobegon" Tests Make It Easy to Cheat

Legitimate "high stakes" standardized tests like the National Assessment of Educational Progress, the College Board, the ACT, and the Armed Services Vocational Aptitude Battery test take great care to see that questions can not be taught. Questions must be equally random to all students taking the test. Accurate assessments can only be ensured by strict security.

School personnel are forbidden from looking at questions that will be used again on the College Board, the National Assessment, or the ACT test, either before or after administration. The test booklets are sealed, either individually, or shrink wrapped in class size packets with clear instructions that the seal only be broken at the time of testing. Tests are delivered in a manner that precludes educators from reading and then teaching the test. Test sessions are randomly monitored to prevent cheating.

Unfortunately, "Lake Wobegon" achievement tests use almost no test security. Teachers receive the test booklets weeks before testing. Test booklets are not sealed. Questions are not rotated. Test proctors are not used. Principals are left in charge of answer sheets. Testing is not monitored by state officials. Answer sheets are not scanned for alterations. Unlike legitimate standardized tests, "Lake Wobegon" tests are administered in ways that make it very easy to cheat.

CHAPTER 5

Deceptive Testing Practices

> "Fraud is infinite in variety; sometimes it is audacious and unblushing; sometimes it pays a sort of homage to virtue."
>
> *Lord MacNaghten: Judgment in Redway vs. Banham, 1986*

Another method of artificially raising test scores is to make certain that slow students will be excluded from testing. However, it is just these youngsters who need extra attention by the school system. Parents need to know if their child is far behind before it is too late.

Many states exclude special education and bilingual students from testing unless their federally required Individual Education Plan (IEP) specifically states they should be tested. Most IEPs simply don't address the issue, so these low achievers are excluded from testing. Only 22 states require special education and bilingual students to be tested unless their IEP specifically states they should not be tested, a crucial difference. Needless to say, all school districts tested their "gifted" special education students.

Excluding these "at risk" students from achievement testing artificially inflates school and district achievement levels. For example, in 1987, Maryland tested 37,000 students at each grade level (33). However, according to the U.S. Department of Education, total enrollment in the third and fifth grade in Maryland is around 44,000 (34). The Maryland Department of Education excluded 20 percent of their total enrollment from testing but regularly issued press releases claiming Maryland was above the national average (10). Similar activity has been reported in New York City (29) and Chicago (30).

Books on Cheating

Another deceptive testing practice involves the use of "test preparation materials

that are often written by the "high stakes" test publishers, such as *CAT Learning Materials* by CTB/McGraw-Hill, and *Scoring High* by Random House. These materials are developed directly from the most recent editions of "high stakes" tests. Many of the review questions in these booklets are identical to the actual questions on "high stakes" tests.

For example, *CAT Learning Materials* preps students on a California Achievement Test thermometer question by telling students how to change a thermometer reading by 10 degrees. One of the questions on the California Achievement Test, Form E, asks students to indicate a thermometer reading 10 degrees higher than the one pictured (35).

Test-Curriculum Alignment

The most common deceptive testing practice is called "test-curriculum alignment." Fifteen years ago, when norm-referenced tests were used solely as instructional aids, "test-curriculum alignment" allowed educators to pick the test that would best aid them in assessing curricular weaknesses.

"Test-curriculum alignment" would be similar to France, Italy and Spain all deciding to test achievement in European history, but France picking a test emphasizing French history, Italy one emphasizing Italian history, and Spain one emphasizing Spanish history. All three would test "above average" in European history.

Now, most states have committees which accomplish such alignment in one of two ways (15,16). Tests are chosen on the basis of how well they match the local curriculum. At first thought this seems like a good idea. However, now that "Lake Wobegon" tests are used to assess educational quality, "test-curriculum alignment" allows administrators to pick tests that guarantee their district will score higher than the national norm group. Remember, the norm group did not get to pick their test, norm groups are selected randomly to represent average national achievement.

After school districts pick a test, they further narrow their curriculum to match the tested content (15,16). Even without frank cheating, such narrowing guarantees that scores will surpass the norm group, because the norm group did not have its curriculum narrowly focused on test content. Most importantly, the curriculum is then actually degraded because it becomes focused on simplistic multiple choice questions (16).

NAEP Testing Does Not Degrade Curriculum

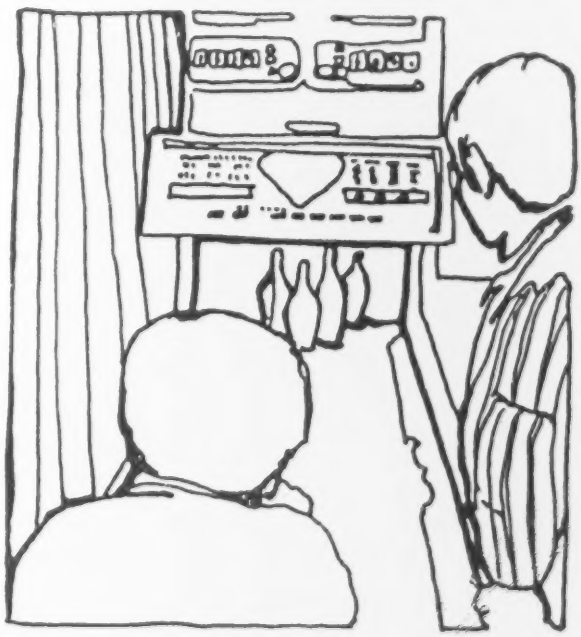
Legitimate "high stakes" standardized testing services never allow their tests or the curriculum to be degraded in this manner. Teaching to legitimate standardized tests is not a problem, because neither the test questions nor the exact test content can be

known in advance. Teaching to a legitimate standardized test, like NAEP, can only be accomplished by broadening the curriculum, not by narrowing it.

For example, the director of the National Assessment of Educational Progress, Mr. Archie Lapointe, was asked by an irate superintendent how he could be expected to do well on NAEP if he couldn't integrate the questions into his curriculum in advance. Mr. Lapointe replied his state would do well on math if the students practiced a lot of math and do well on reading if they read a wide variety of literature.

Mr. Lapointe then provided a copy of the broad range of skills that NAEP will cover. The superintendent, who wanted specific test content, was not satisfied, so Mr. Lapointe added that the math section would have a lot of numbers and the English section would have a lot of letters!

This apocryphal story underlines a crucial point: **the curriculum will be degraded when tests are "high stakes," and when specific test content is known in advance**. NAEP will soon become a "high stakes" test because of state-by-state comparisons, but it will broaden curriculum, not narrow it, if test content is kept broad and if test questions remain secure.



CHAPTER 6

Misleading Reporting Methods

"There are three kinds of lies: lies, damned lies, and statistics."

Unknown

School administrators have used statistics to change "below average" scores into "above average" scores. For example, Hawaii is one of only three states to report with stanines, which are a statistical groupings of scores into nine categories. Students are grouped into stanines one to three (below average), stanines four to six (average), and stanines seven to nine (above average). However, students in stanine four, and half the students in stanine five, are actually performing below the 50th percentile.

Hawaiian officials reported that "77 percent of Hawaii's third graders were average or above average in reading" (stanine four to nine), instead of reporting that the majority of Hawaii's third graders were testing below the national average in reading (national percentile rank of 42) [see Appendix 1].

Pick Your Scores

We also found that districts often choose percentile ranks instead of normal curve equivalents as required by the federally funded Chapter I program. A mean normal curve equivalent of 60 for a district can be reported as a mean individual percentile rank of 70, which looks better, or a mean group percentile rank of 90, which looks even better. Group percentile ranks are particularly deceptive because they are interpreted as the average student's score when, in fact, they are group comparisons, not average group scores.

Interestingly, we did not find any districts using group percentile ranks if they tested below the national norm of 50. Group percentiles make low scoring districts appear even lower, for the same statistical reason that they make higher scoring districts appear even higher. Most "below the norm" districts reported with

normal curve equivalents, the reporting method that makes their scores the closest to the 50th percentile.

Group percentile ranks, stanines, and grade equivalents are some of the more misleading statistical methods that school officials use to report "high stakes" test scores to the general public. However, districts can also buy "low socioeconomic norms" and "large city norms" (36), both of which will make scores appear higher. These norms allow a low scoring district to compare itself with a low socioeconomic or inner-city group of children, instead of the national population. Such comparisons are legitimate when the tests are used for instructional purposes in low socioeconomic districts, but not when they are used for "accountability."

National Average Vs. National Norm

The words "average" and "norm" are used interchangeably by the public. However, they have very different statistical meanings. A school district could be "above the national norm" and still have the lowest achievement scores in the nation. For example, in 1987, West Virginia had the lowest average CTBS math scores of any third grade tested statewide in the USA (10). However, West Virginia school officials were able to claim that West Virginia's third graders were above the national norm with a 54 percentile (meaning they did better than 54 percent of the publisher's 1981 "national norm group").

The student achievement report forms that public educators use to report student achievement to parents also imply that "above the national norm" is the same as above the national average. For example, the 1988 SRA Narrative Report for explaining norm-referenced test scores to parents explains how their child's "scores fall in relation to below average, average, or above average," and then lists normed-referenced test scores (37).

Five states use the California Achievement Test parent report (I-131-PR), which tells parents: "The national average for each test would be considered a percentile rank of 50" (38). In fact, the 50th percentile is the average score of the norm group tested "cold" in 1985, a norm group that included special education students.

Most state publications imply that "above the national norm" means "above the national average." Delaware's publication (39) states that "Average Delaware student performance was above the national average at all grade levels on the Total Basic Battery except for grade 1. Average scores higher than 50 are above the national norm." Many states had impressive graphs with their achievement towering above national achievement. However, few state publications honestly described the difference between norm and average.

State Education Department press releases also imply that average and norm are the same. A 1987 West Virginia "Quick Facts" press release displayed grade-by-grade, above-average test scores with an explanation that "national average equals 50" (40). A June 16, 1987, New Mexico press release stated scores "are

above the national average, which is the 50th percentile" (41). In fact, the 50th percentile in this case is the average score of the norm group tested "cold" in 1981.

As can be seen, the states use norm-referenced scores to issue glowing reports on state achievement. Even the few states with "below average" scores eventually issue publications that emphasis steady, usually dramatic, gains in achievement. Educators claim the tests are used for "instructional" purposes, but it is hard to guess that from their glossy brochures, rosy press releases, and optimistic "accountability" reports.

CHAPTER 7

Suspicious Patterns In Test Scores

"Wisdom liketh not chance."

English Proverb

Some states have very "high stakes" tests without any serious security regulations. Arkansas, Georgia, Kentucky, Oklahoma, South Carolina, Tennessee, and West Virginia all have "high stakes" testing programs that have administered the same test form year after year. We have received allegations of cheating from most states but especially these seven states. None have significant test security measures. However, Arkansas, Oklahoma, and Kentucky have recently taken some steps to improve test security.

In these seven states, we see certain patterns in test scores not evident in states with adequate security measures. We feel these patterns, especially when they occur in combination, may be evidence that testing irregularities are occurring.

The Hi-Lo Show

In these seven states, norm-referenced test scores are much higher than any other indicators which often correlate with school achievement. Appendix I lists the most recent "Lake Wobegon" test scores for these eight states.

Compare West Virginia's test scores from Appendix I with the other indicators listed in Appendix III. In the most recent scores on the Comprehensive Test of Basic Skills (CTBS), 72 percent of West Virginia's third graders, 64 percent of its sixth graders, 52 percent of its ninth graders, and 58 percent of its 11th graders tested above the national average. In contrast, Appendix III shows that West Virginia has the fourth lowest college entrance scores (ACT), the ninth lowest ASVAB scores in the nation, and they rank 39th in childhood poverty.

Likewise, Arkansas is above the national norm on the Metropolitan Achievement Tests. Appendix III shows that Arkansas ranks 46th out of the 50 states on

college entrance scores, 38th on the Armed Services Vocational Aptitude Battery, and 47th on childhood poverty.

These additional indicators will provide a general idea of achievement in your school district. Although all of them have significant limitations, they will let you compare different methods of measuring achievement. If your norm-referenced scores are much higher than the other indicators, be alert to the possibility that cheating may be occurring in your schools.

Downhill in Secondary School

These seven states generally have much higher elementary scores than secondary scores. Most states with low stakes testing, as well as those states with the better security measures, do not show this marked difference between elementary and secondary achievement.

Remember, most, but not all, cheating occurs at the elementary level because at the present time, norm-referenced tests are the only available method of comparing local elementary achievement to national elementary achievement. Secondary schools have many methods to obtain national comparisons such as college entrance scores, ASVAB scores and graduation rates.

American elementary schools rely exclusively on "Lake Wobegon" tests or on state developed criterion-referenced tests (which have similar problems with test security, test teaching and the inevitable curriculum degradation). Such tests have assumed exaggerated importance in American elementary schools because the public's focus on test scores, when combined with lax test security, makes cheating the easiest way to claim that achievement is improving.

Is it possible that the elementary schools in these seven states are just doing better than their secondary schools? If that were true, why hasn't the increased student achievement eventually shown up in their secondary schools?

Language and Math Easiest for Cheating

In these seven states, the subtests which are easier to cheat on, such as language and math, are generally higher than reading scores. For example, Tennessee's elementary math scores are generally ten points higher than their reading scores. The difference then decreases in secondary grades.

The math and language subsections of "Lake Wobegon" tests are easiest to cheat on because of the nature of the questions. The questions are straightforward, and the subjects lend themselves to being easily taught by teachers who remember test items. However, the reading subsection consists of reading passages that are followed by three to eight comprehension questions. Cheating on the reading subtest is possible, but it is much more difficult.

Instructional differences can explain some variation between reading and math scores, but it is unlikely that it explains all the differences. Math, in particular, is more responsive to a curriculum which narrows itself to certain particular skills, like long division. However, we see little subtest variations in states with good test security procedures.

Many other states have one or two of these suspicious score patterns, but these seven states have all three. It is not possible to simply look at a state or district's test results and say with certainty that cheating is occurring. These are simply three clues, some red flags, which should alert responsible officials to that possibility.

Is Cheating Occurring in Your District?

Based on findings in our latest survey, we propose three easy methods which state legislators and school board members can use to evaluate local "high stakes" test scores. If state legislators and school board members choose to implement "high stakes" tests, they should understand test patterns, usefulness, and limitations. **Remember, these methods do not prove cheating is occurring; they should instead serve as a basis for further questions.**

Are the district scores in line with other indicators of achievement? As explained in Appendix III, local districts can easily and inexpensively obtain other indicators of local achievement such as college entrance scores (ACT for 28 states and SAT for 22 states) and ASVAB scores.

Are relative achievement levels similar for all grade levels? Scores should be at similar national percentile ranks for all grade levels tested. Socioeconomic factors are easily the strongest determinants of school achievement, and those factors are present at all grade levels. A steep drop in secondary scores may mean that cheating is occurring in your elementary schools.

Are your language and math scores close to your reading scores? If your reading scores are much lower than the language and math scores, cheating may explain the difference. Instructional and curricular differences may also explain some of these variations.

These are simply three patterns we have observed in our study. They are meant to serve as a basis for further exploration, not as proof that cheating is occurring.

CHAPTER 8

Steps To Avoid Cheating

> "As new cases occur the law is perpetually found deficient. It is therefore perpetually necessary to make new laws."
>
> *William Goodwin: An Inquiry Concerning Political Justice, 1793*

If state legislators and school boards decide to test, they should first decide if they want tests for accountability purposes or for instructional aids. If it is the latter, then scores should not be compiled for classes, schools, districts, or states. If they are compiled, they will eventually be made public by "freedom of information" requests filed by newspapers, public interest groups, and concerned citizens.

It must be remembered that school administrators directly benefit from inflated test scores, and some administrators may not be inclined to institute effective test security. It might be wise to separate instruction from instructional assessment by contracting with a reputable outside testing service to design, administer, and score your district's achievement tests.

However, state assessment, research, and accountability departments can provide accurate student achievement data if they have the necessary budget, and if they are adequately insulated from departmental political pressure. **If budgets are tight then decrease the amount of testing, but don't test extensively unless the assessment department has the resources and the authority to ensure test security.** And, if assessment directors are hired and fired by the superintendent, don't expect achievement data to be free from politics.

If you decide to administer "high stakes" tests, make certain that you have clear "high stakes" security policies in place. These policies should be state policies and they should apply to any group administered test used in the state that reports class, building or district achievement, including tests administered independently in local school districts. Educators deserve to be told unequivocally what is and is not ethical. General, loosely worded statements in a "Code of Testing

Ethics" should be avoided. Educators should be clearly told what test practices are unacceptable, and such practices should be made illegal.

If test scores are compiled for classes, schools, and districts, then the tests are "high stakes" and will be used for accountability purposes by the public and the press. A number of security precautions then should then be taken by school boards and state legislatures to prevent testing irregularities, misleading testing practices, and deceptive reporting methods. **Don't administer "high stakes" tests without "high stakes" security.**

Suggested Steps to Limit Cheating

1] **States and school districts should adopt policies which clearly forbid school personnel to look at "high stakes" test questions except as needed during administration.** These instructions should be contained in both state and local school testing policies and should be prominently printed by the publishers on test booklets. Teachers and principals should receive clear and unambiguous messages that looking at test booklets is illegal and unethical.

2] **Testing should be done in the fall, as early as possible, preferably in September.** Thus, students will be tested on material they learned the previous year and during the summer. It will reduce the valuable class time that is spent prepping for the test, and reduce teachers' motivation to cheat because teachers will no longer see the test as assessing their teaching skills.

However, fall testing does not reduce administrators' incentives to cheat and thus does not obviate the need for additional test security measures. In addition fall testing measures summer learning as well as last year's instruction. Therefore, fall testing is not as "pure" a measure of instructional effectiveness as spring testing.

3] **"High stakes" test booklets should be sealed individually and then shrink wrapped in class-size packets.** This will cost a little more every year, but is crucial for "high stakes" test security. Instructions on the test should clearly state that only the students can break the seal and that school personnel are not to look at the tests.

4] **Tests should not be given to teachers until the day of testing and should be turned back in after that day's testing is completed.** The publishers should clearly state this in their examiner's manual. Principals should receive the box of sealed test booklets the day before testing. Tests should be picked up the day testing is completed. UPS and other such companies could be contracted to deliver the test the day before testing.

5] **States should have test monitoring teams.** They should be empowered to walk unannounced into schools to inspect test administration and test

security. State legislators should see that state testing coordinators have the personnel, financial resources, and political authority to effectively carry out such monitoring. Confirmed irregular testing practices discovered in schools should be reported to the press.

6] **Volunteer test proctors should be requested from the community if resources are not available to hire a reputable testing firm to administer the tests.** In any case, teachers should not be allowed to test their own classes. It would be a simple matter to require that teachers test classes other than their own, preferably one at a different grade level.

7] **All children, including special education and bilingual students, should be tested and their scores reported with their current grade level.** Special education and bilingual students should be specifically included unless their Individual Educational Plan (IEP) excludes them for sound reasons, not just because of expected low scores. The publisher's manual should clearly state that special education and bilingual students are to be tested. State and district reports should clearly state the number of children excluded from testing and the the reasons for such exclusions.

8] **Equivalent "sister" tests should be randomly substituted in those school districts with suspiciously high test scores.** Most commercial tests have equivalent forms of the same test for sale. They test exactly the same content but use different test questions. Classes that have been unethically prepped will show a significant drop in scores when the same curricular content is tested with an equivalent form of the same test. School personnel should be warned that such substitutions will occur.

9] **New tests forms should be purchased every year by those districts that can afford it.** Under no circumstances should tests be used for more than two years, because the questions will eventually become known by the educators. When districts have to reuse a test, the initial purchase order should be divided to include as many equivalent test forms as the publisher sells. For example, if the same test is to be reused, then one-half of the new tests should be Form A and one-half Form B. The tests should then be rotated in such a way that teachers will be using a different test form every year.

Matrix sampling, as used by the state departments of education in California, Massachusetts, and Maine also effectively addresses the problem of test question familiarity by administering multiple test forms every year. However, matrix sampling will not give individual student's scores, but individual student scores could be obtained from an additional "low stakes" test; that is, a norm-referenced test that only reports student scores, not class, building, district, or state scores.

10] **Answer sheets should be scanned by the company that scores the test, both for suspicious erasures and for cluster variance (suspicious class groupings of similar responses).** Such technology is readily available and inexpensive, although not free.

11] **Scores should be reported in normal curve equivalents using the most current national norms available.** For statistical reasons, normal curve equivalents are the reporting method which is least sensitive to small year-to-year changes in achievement. As such, they tend to deemphasize test scores and help prevent the inevitable degradation of curricular content that occurs when scores on multiple choice tests are emphasized.

Eventually, the publishers should be required to develop scaled scores anchored on past national achievement. This would allow the public to see changes in achievement as well as current state and national comparisons. Stanines, grade equivalents, and group percentile ranks are deceptive and should not be used to report achievement scores to the public.

12] **Current annual norms calculated from a representative national sample of current users should be used in all your schools, and all publishers doing business in any school district in your state should be required to sell them instead of old norms.** Annual norms will more closely approximate a current year average, which is the meaning the public infers from the phrase, "above the norm." Annual norms will give school boards and state legislatures a better idea of how local schools compare to the current national average.

However, annual norms present educators with a "moving target," and districts would have to surpass that year's national achievement gains in order to show progress. In addition, annual norms should be calculated from a representative national sample of users of the test, not from all users. Otherwise, tests, such as the California Achievement Test that are used in a disproportionate number of low achieving districts, will have deceptively easy norms.

13] **Parent report forms should be required to clearly state the limitations of norm-referenced tests.** The norm group should be described including the year the norm group was tested. The percentage of special education students tested in the norm group should also be clearly stated. Parents should be told that "above the norm" scores does not mean their child is currently achieving above the national average.

14] **Test publishers and scoring companies should be required to report achievement results yearly to the state department of education for all of the publisher's tests used in the state.** Districts should also be required to report their test scores to the state department of education. This would allow publishers, state departments of education, citizen groups, and the

federal government to more easily monitor test accuracy.

15] **"Test-curriculum alignment" committees should be disbanded.** Only general information about test content should be used when picking a "high stakes" test. The actual test should not be made available by publishers until shortly before testing. Even "sister" forms of the test, if studied in advance of testing by school personnel, will result in "test-curriculum alignment" and inevitably, in degradation of your curriculum.

16] **Writing tests should be administered along with multiple choice tests.** Writing tests promote the development of critical thinking skills and they also help deemphasize the importance of multiple choice tests.

17] **Test preparation books, such as *Scoring High* by Random House, and *CAT Learning Materials* by CTB/McGraw Hill should not be allowed in your schools.** These materials are modeled directly from "high stakes" test questions and their use destroys the validity of the needed inferences from test scores to the larger domain of overall achievement. Any test publisher that sells such books, or that allows their test questions to be used in such books, should be banned from doing business in your state.

18] **Any business, test publisher or individual, who sells "high stakes" group achievement tests in your state should be liable for punitive damages for:**

a. selling tests whose norms are more than two years old,
b. selling tests whose norm sample has not been approved by your state department of education,
c. selling tests that are unsealed,
d. selling tests that do not clearly forbid teachers from reading the test except as needed for administration of the "listening" subtest,
e. selling tests whose norms tables do not report with normal curve equivalents,
f. selling tests whose administration manual does not specify the above security measures, and
g. willfully withholding evidence of testing irregularities from responsible state officials.

NAEP Participation Opportunity

State and local officials should insist that their state participate in the upcoming state-by-state evaluation by the National Assessment of Educational Progress (NAEP). This testing will be conducted first in 1990 (eighth grade math) and again in 1992 (eighth grade math, and fourth grade math and english) under the supervision of the U.S. Department of Education. Test security will be under the control of

personnel from the Educational Testing Service, not local school officials. At this time eleven states, listed in Appendix I, have decided not to participate in NAEP.

State and local political leaders should also lobby their federal representatives to change the federal law that prohibits NAEP evaluations of local district achievement. Presently, federal law specifically prohibits NAEP from being used to assess local district achievement, even if the district is willing to pay the extra cost of administering it to their entire school district.

CHAPTER 9

The Effects of Inflated Test Scores

"A whale ship was my Yale College and my Harvard."

Herman Melville: Moby Dick, 1851

The accusation of widespread cheating by public school officials is a serious one. It is not that educators are any less ethical than other professionals. Let's not pretend that some lawyers would refuse the opportunity to cheat on the bar, if they could obtain the exam a week before testing, or some physicians, if they could obtain copies of their medical boards in advance. A closer analogy might be if state legislatures decided to evaluate physicians on the basis of a general health information quiz taken by their patients. The physicians would certainly teach those patients the exact quiz questions at each office call, if the state let physicians preview the patient's test in advance.

"Lake Wobegon" Tests not Designed for Accountability

The norm-referenced tests now used to assess public educators were never designed for such a task. They have evolved from being instructional and curricular aids into instruments of public accountability because of external political "accountability" pressures.

Twenty years ago norm-referenced achievement tests were mainly used for instructional purposes. Teachers used them to determine which students were behind and if the class needed more work in one subject than another. Class scores, school scores, district, and state scores were either not compiled or not made public. Norm-referenced tests were used to help children, not to evaluate educators.

However, that changed when the press, school boards, state legislatures, and school reform groups, like Friends for Education, started insisting on accountability. Almost overnight, the tests were asked to serve an accountability purpose

instead of just an instructional one. They have since become the principle local yardsticks of American educational progress.

They will remain tools of accountability as long as the press insists on printing test scores, as long as school boards and state legislatures insist on knowing student achievement, and as long as the public insists on accountability. It seems unlikely they will ever again be solely instructional aids. Therefore, the tests must be modified to serve their present function.

The glowing press releases, glossy student achievement brochures, "good news" parent report forms, and optimistic official "accountability" reports put out by American public schools are testimony to the fact that public educators themselves now use norm-referenced achievement tests to measure school quality. And, for the last 15 years, American educators have found it easier to improve test scores than to improve public schools.

Racists Tests

These commercial tests are also used in a racist manner. It is not racially biased questions that damage black children now; such questions have been largely eliminated by the publishers. It is the appearance of high scores that is racist. Improving scores convinces the local press, parents, and politicians that things are getting better in Chicago. Inflated test scores are the "good news" that effectively resists the reforms needed in America's inner-city schools.

What would be the effect on Chicago schools if they had to report that Chicago's children only performed at the 20th percentile, not the 50th? What would happen if politicians in America's inner-city schools had to report that no improvement has occurred in their schools in the last decade? Perhaps substantive improvements would finally be made.

State legislatures and school boards need accurate measurements of local achievement. Local officials can not operate blindly, they need to know what children know and when they know it. How can local officials reform American schools when their principle yardsticks tell them they already have?

All American schools, not just white middle class ones, must insist that children achieve what is expected before they are promoted. Otherwise, they become hopelessly lost after being socially promoted into grade levels where they can not compete, where their self-esteem is shattered, where they become depressed, or delinquent, or pregnant; and where they eventually drop out.

All American children, regardless of social class, deserve schools with both high expectations and high standards. Elite private schools clearly tell children and parents what is expected at each grade level, and they are expected to achieve it before they are promoted. We need the same love and respect shown for all children.

APPENDIX I

NATIONALLY NORMED TEST RESULTS OF THE 50 STATES

Methods

Friends for Education surveyed the 50 state departments of education in the spring of 1989. We sent registered letters to the superintendents of public instruction in all 50 states requesting them to supply us with the names of all tests administered on a statewide basis in their state.

We requested the dates of administration, norms used, statistical method used to report the scores, number of students tested, the latest reading, language, math and basic battery scores, percentage of students testing at or above the 50th percentile on their aggregate total battery score at each grade level tested, and the percentage of that state's school districts testing at or above the 50th percentile on their aggregate total battery score.

We also requested copies of the test security and procedure policies for all tests, including any nationally normed, criterion referenced and graduation exams, the state departments of education administers on a statewide basis. We also requested copies of any state rules and regulations for the norm-referenced tests administered independently by their local school districts.

The state scores were obtained directly from state departments' of education publications, and the latest available scores for nationally normed tests as used in all 50 states are listed at the end of this Appendix. We attempted to obtain similar score reporting methods for all states, but many states used alternative statistical reporting methods. National mean individual percentile ranks were obtained whenever possible.

The percentage of students and the percentage of districts testing at or above the national norm (% students and % districts > or = 50) was calculated from the composite or total basic battery scores when available. When composite or total basic battery scores were not available (N/A), we used language, then math, then reading scores. Unless otherwise stated, the "national norm" is 50.

Ten states denied knowledge of individual district results, so Friends for Education telephoned the largest school districts in these 10 states. We asked if the "majority of their total basic battery or composite scores were above or below the national norm of 50." If some grade levels tested below and others tested above, then "above 50" was defined as an average normal curve equivalent for grades one through six that exceeded fifty, or a majority of elementary grades that exceeded 50 on the composite or total basic skills score. If the total basic skills or composite scores

were not available, then scores for all available subtests were used.

In addition to the latest scores, we include narrative details of the test security practices for each of the fifty states. These policies are presented in table form in Appendix II, and detail the test security policies for any nationally normed test administered in the state. If the state does not administer any nationally normed test, then the security policies for criterion referenced tests are listed. Four states administer no statewide test of any kind, nor do they collect data from the local districts in their states. For these four states the table lists state security regulations for the norm-referenced tests administered locally.

We requested that states respond to the following survey:

Survey of Testing Policies and Practices

Please check all that apply:

[1] Does «state» have any state testing program?

a. yes, off-the-shelf, norm-referenced achievement test
b. yes, customized, norm-referenced achievement test
c. yes, locally developed test, normed by equating study
d. yes, locally developed test, not nationally normed
e. yes, criterion referenced test, not nationally normed
f. yes, minimum competency test
g. no, and we are not planning one
h. no, but we are planning one (circle which kind)
i. we collect local district's test results
j. other (please specify)

[2] Does «state» have a state policy on test security? Please include a copy of your policy.

a. yes, for both state and district testing programs
b. yes, but only for tests the state requires
c. no, we test but we have no state policy
d. we don't test, so such policy is up to the districts
e. we are developing such a policy
f. other (please explain)

[3] Does «state» have a policy on testing procedures? Please include a copy of your policy.

a. yes, for both state and district testing programs
b. yes, but only for tests the state requires
c. no, we test but have no written testing procedures
d. we don't test, so such policy is up to the districts
e. we are developing such a policy
f. other (please explain)

[4] Does «state» investigate allegations of cheating or irregular testing practices? Please include a copy of your policy for such investigations.

a. yes, a formal state investigation takes place
b. yes, informal state investigation (e.g. phone calls)
c. such investigations are referred to the district
d. we don't test so it is up to the districts
e. other (please explain)

[5] How many such allegations have you investigated over the last five years?

a. 0
b. 1-2
c. 3-5
d. 6-10
e. 11-20
f. more than 20

[6] How many cases of test security violations has «state» confirmed over the last five years?

[7] Does «state» allow teachers to see actual questions from nationally normed achievement tests, either before or after testing? If no, please include a copy of your policy forbidding teachers to look at the tests.

a. yes, we encourage them to read over test questions
b. yes, but we don't encourage it
c. yes, the publisher's examiners manual recommends it
d. no, we expressly forbid them to look at the test
e. we have no state policy, it is up to the districts
f. we don't test so such policy is up to the districts
g. other (please explain)

[8] Please indicate if state policy specifically excludes any of the following groups of students from testing? Please include your test exclusion policy.

a. behavioral disordered
b. learning disabled
c. limited english proficient students
d. mentally impaired
f. visually handicapped
g. hearing handicapped
h. emotionally impaired
i. we test all students that can hold a pencil
j. we have no state policy, it is left to the districts
k. we don't test so such policy is up to the districts
l. other (please explain)

[9] Does «state» use any technological devices to detect cheating?

a. yes, optical scanning to detect suspicious erasures.
b. yes, computerized variance analysis of answer sheets
c. no, no such devices are used

d. we don't test
e. other (please explain)

[10] Please check those quality controls that «state» uses to discourage cheating.

a. seals on test booklets
b. yearly rotation of all test questions
c. yearly rotation of some test questions
d. matrix sampling test with multiple forms
e. random spot substitution of equivalent tests
f. tests are administered in early fall
g. tests are delivered to schools the day before testing
h. tests are handed to teachers the day of testing
i. teachers do not administer tests to their own class
j. teachers are clearly told not to look at the test
k. outside test proctors are used
l. we don't use any quality controls
m. we don't test
n. other (please explain)

Thirty-eight states responded within one month, and the remaining states responded to a follow-up telephone inquiry. On August 5, 1989, the data sheets for each state were sent back to their state testing coordinators by registered mail for verification and corrections. The states followed by an asterisk (*) responded and verified our data, although some objected to our wording.

Results

We found twenty-seven states have tested statewide with a commercially available, "off the shelf," norm-referenced test within the last three years. Some states administer the test to all students at all grade levels every year, such as Arizona and Kentucky, while other states test a stratified random sample of students at a few grade levels every three years, such as Utah and Wisconsin.

Eight states administer a locally developed test on a statewide basis that has been nationally normed in some manner in the last four years by equating it to a commercially available test. Many of these states report with scaled scores that only emphasize the state average, such as California, Maine, Massachusetts, Pennsylvania, and Texas, but their tests have been equated to a nationally normed test in the recent past.

Fourteen states have not administered any nationally normed tests on a statewide basis in the last three years. However, six of these states administer a statewide criterion referenced test. In all these 14 states, local school districts administer norm-referenced tests independently, that is, without statewide control. Four of these 14 states compiled local districts' norm-referenced test scores and the

results the state compiled are so listed. In ten of these 14 states, the state denied knowledge of individual district results. The largest districts in these ten states were then surveyed by telephone as explained above, and our findings are listed.

Forty states have reported statewide "nationally normed" achievement scores (various grades between one and twelve) in the last three years. These forty states include states that use an "off the shelf" norm-referenced test (28 states), states that use a locally developed test that has been equated in some manner with a nationally normed test (8 states), and states which collect local district test data, but who do not administer a statewide test through their state department of education (4 states).

The states reported test results by many different statistical reporting methods, such as individual percentile ranks, group percentile ranks, normal curve equivalents, scaled scores, stanines, and grade equivalent scores. Friends for Education attempted to obtain mean individual national percentiles from all states but that proved impossible due to local variations in reporting methods. We reprinted the scores as reported in the state's publication. We did not perform any additional statistical operations on the data, except for the two states that will only release data in stanines. In those two states, we listed the stanine scores and made conservative estimates of the national percentile rank from stanine data.

We found the states use the following tests:

IOWA TEST OF BASIC SKILLS:
Arizona
Colorado
Georgia
Idaho
Missouri
Virginia

STANFORD ACHIEVEMENT TEST:
Alabama
Hawaii
Mississippi
Nevada
South Dakota
Tennessee

METROPOLITAN ACHIEVEMENT TEST:
Arkansas
Oklahoma
Rhode Island
Washington

LOCALLY DEVELOPED TEST:
California
Connecticut
Illinois
Massachusetts
Maine
Pennsylvania
Texas

COMPREHENSIVE TEST OF BASIC SKILLS:
Kentucky
New Mexico
South Carolina
Utah
West Virginia
Wisconsin

CALIFORNIA ACHIEVEMENT TEST:
Indiana
Louisiana
Maryland
New Hampshire

STATES WHERE DISTRICTS TEST INDEPENDENTLY:

Alaska	Michigan	Nebraska	North Dakota
Iowa	Minnesota	New Jersey	Ohio
Kansas	Montana	New York	Oregon
			Wyoming

Conclusions

We conclude that forty-eight of the fifty states are still scoring above the "national norm" on "standardized, nationally normed" achievement tests two years after our original "Lake Wobegon" study. Only Arizona and Louisiana had a majority of their scores below the publisher's national norm. A few states had one or two grade levels or an occasional subject that was below average, but most states had all subject areas and all grade levels well above the publisher's "national norm."

Ninety-five percent (38 of 40) of the states that test statewide or that collect local district results reported that the majority of their elementary scores were above 50th percentile. Eighty-eight percent (35 of 40) of these states had all their elementary scores, in all subject areas and all grade levels, above the publisher's national norm of 50.

Thirty-nine states have reported "nationally normed" scores at secondary grade levels (various grades between 7 and 12). Ninety-five percent (37 of 39) of these states had the majority of their secondary scores above the 50th percentile. Ten of these 37 states had one or more subtests, usually reading, below the 50th percentile, but the majority of their secondary scores were all above the 50th percentile.

Many of these states rank below the national average on standard barometers of excellence, such as childhood poverty, college entrance scores, and Armed Services Vocational Aptitude scores (Appendix III). Such states include Alabama, Arkansas, Georgia, Kentucky, Mississippi, Missouri, New Mexico, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, and West Virginia—all of which are testing above the publisher's "national norm" on one of the norm-referenced tests.

The majority of students in all these states are told they are above average. For example, 65 percent of Georgia's second graders tested above the publisher's national norm on the total battery of the Iowa Test of Basic Skills, and 75 percent of Kentucky's third graders tested above the norm on the Comprehensive Test of Basic Skills in spite of the fact that Georgia and Kentucky have among the lowest literacy

rates, lowest college entrance scores, and lowest Armed Services Vocational Aptitude Battery scores in the nation.

All the fourteen states where the districts test independently had a majority of their scores above the 50th percentile. The ten states where we surveyed the largest districts (the states where the State Departments of Education had no knowledge of individual district results) had a majority of surveyed districts above the 50th percentile. As with the statewide results, elementary grades in these 10 states were more likely to be above 50 than were secondary grades. These states are all northern and western states that generally rank above the national average on standard barometers of excellence.

Friends for Education concludes, after sampling their largest districts, that virtually all the districts in these 14 states are testing above the publisher's "national norm" of 50. If poor southern states have a majority of districts above average, then wealthier northern and western states should have all but their inner-city districts above average on the same tests. For example, in Rhode Island, all but two inner-city school districts tested above the publisher's norm and all of Maryland's 24 districts (except Baltimore City) were above the norm.

These larger districts often include metropolitan areas with many inner-city children, a group often thought to be below the norm. Instead, most inner-city districts were claiming to be above average at elementary level, including Trenton and East Orange, NJ; Kansas City, KS; Omaha, NE; Columbus, Toledo, and Akron, OH; and Grand Rapids, MI.

Friends for Education obtained results from a total of 5,143 elementary districts nationwide; 4,192 or 83 percent had the majority of their elementary scores above the publisher's "national norm." We obtained results from a total of 4,501 secondary school districts nationwide and found that 3,264 or 73 percent were scoring above the publisher's "national norm." 4,948 of these district scores were obtained directly from the state department's of education publications, and 95 were obtained from our telephone interviews.

Many of these districts were from states that rank low on other barometers of student achievement. For example, the scores we obtained from the 5,143 elementary districts include all the school districts in the south, except for Texas. All the states in the south, except Louisiana, reported that the majority of their local districts were achieving "above the national norm."

For example: 83 percent of Alabama's 128 districts tested above the national norm on first grade tests; 93 percent of Georgia's 186 districts scored above average on 2nd grade tests; 89 percent of South Carolina's 92 districts scored above the norm on 4th grade tests; 99 percent of Tennessee's 139 districts scored above the national average on their second grade exam; and 99 percent of Kentucky's 179 school districts and 100 percent of West Virginia's 55 districts tested above the national norm at the 3rd grade.

Because our district data includes an over-representation of low socio-economic districts in southern states, we estimate that nationwide, 90 percent of

American school districts are testing "above the national norm" at elementary level and more than 80 percent at secondary level.

State-by-state comparisons with these scores are of no value because different tests, different norms, and different reporting methods are used. CAT, SRA, and ITBS had the highest scores nationwide, but it is important to look at the states using the tests. Because of their high college entrance and ASVAB scores, and their low percent of childhood poverty (Appendix III), Maryland and New Hampshire would be expected to do well on the CAT (Maryland uses the older test), but North Carolina's scores are unexpectedly high. Iowa, Idaho, and North Dakota would likewise be expected to do well on ITBS, but the scores in Missouri and Georgia are higher than expected. Oklahoma and Arkansas have unbelievably high scores on the MAT, a test found to be used, together with CAT and CTBS, in many southern and inner-city districts.

We conclude that "nationally normed" testing, as conducted by most states is deceptive. "Lake Wobegon" testing allows most children and virtually all local school districts to test "above the national norm." "Lake Wobegon" tests let parents, the press, the board of education, and the state legislatures, believe that their state is not part of "A Nation at Risk."

NATIONALLY NORMED TEST RESULTS OF THE 50 STATES

States followed by an asterisk (*) responded to our registered letter and certified their state data is correct. States without an asterisk did not respond, but the data listed was obtained directly from their state publications.

Unless otherwise stated, the national norm is 50.

ALABAMA* *APRIL 1989* THE STANFORD ACHIEVEMENT TEST FORM F 1982 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50*	% DISTRICTS > OR = 50
1	55,583	57	N/A	56	57	57%	106/128 (83%)
2	53,454	52	N/A	65	58	56%	111/128 (87%)
4	52,159	50	58	59	54	55%	98/128 (77%)
5	51,244	48	55	60	53	53%	80/128 (63%)
7	49,340	43	53	53	50	51%	58/128 (45%)
8	47,454	52	55	54	53	54%	77/128 (60%)
10	43,029	47	55	54	51	52%	72/127 (56%)

Reporting method: mean national individual percentiles

Source: Alabama Chief State School Officer's Report

* estimated from stanines

TEST SECURITY IN ALABAMA

Besides the Stanford Achievement Test, Alabama also administers a criterion referenced test and a graduation exam. The new, secure version of the Stanford Achievement Test will be administered in the 1989-1990 school year. Alabama is in the process of further improving their test security guidelines. Alabama rotates questions every year on their graduation exam.

New regulations for the Stanford prohibit teachers and administrators from looking at test booklets, require that tests be given to teachers on the day of testing, require random monitoring of test administration and require special education students to be tested unless their IEP prohibits it.

However, the Stanford booklets are not sealed, fall testing is not required, outside test proctors are not required, questions are not rotated every year, and answer sheets are not scanned for suspicious erasures or routinely analyzed for cluster variance.

Alabama is currently one of only three states that have regulations governing the administration of the commercial norm-referenced tests selected and administered by local school districts.

ALASKA* *1987-88* VARIOUS NATIONALLY NORMED TESTS ITBS AND SRA MOST COMMON

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50*	% DISTRICTS > OR = 50*
K-3	N/A	51	56	61	N/A	53%	30/55(55%)
4-6	N/A	52	56	56	N/A	58%	23/55(42%)
7-8	N/A	53	57	57	N/A	59%	25/54(46%)
9-12	N/A	46	54	57	N/A	55%	22/54(41%)

Reporting method: mean individual percentile ranks
Source: Basic Skills Performance of Alaska's Students
* in language

TEST SECURITY IN ALASKA

Alaska does not administer a statewide nationally normed achievement test. However, all the local districts test with various commercial tests, and the ab ve scores were compiled and published by the Alaska State Department of Education. The Alaska Board of Education recently required statewide achievement testing which will begin in 1989 with the administration of the 1985 Iowa Test of Basic Skills. The law requires that the state use the same test for six years.

The Alaska Department of Education plans a fall administration to test special education students, unless their IEP excludes them.

However, there are no state regulations that forbid teachers from reading the test, that require teachers to obtain the test booklets no earlier than the day of testing, that require the booklets to be sealed, or that require the testing to be monitored by state officials. In addition, there are no regulations that require outside test proctors, that require questions to be rotated every year, or that require answer sheets to be routinely scanned for suspicious erasures or analyzed for cluster variance.

Alaska is one of the eleven states that is not planning to participate in the 1990 National Assessment of Educational Progress state-by-state evaluation.

ARIZONA* *APRIL 1989*

THE IOWA TEST OF BASIC SKILLS
FORM J 1988 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
1	19,034	43	58	49	51	N/A	N/A
2	46,566	49	57	56	55	52%	80/189(42%)
3	46,513	46	51	43	46	44%	55/189(29%)
4	44,835	47	47	43	47	47%	34/187(18%)
5	43,126	49	50	45	49	49%	63/187(34%)
6	42,423	48	49	47	48	48%	58/187(31%)
7	39,302	51	52	47	51	50%	79/182(43%)
8	37,878	49	55	46	50	52%	107/179(60%)

TEST OF ACHIEVEMENT AND PROFICIENCY
FORM J 1988 NATIONAL NORMS

9	37,873	54	52	45	49	49%	44/101(44%)
10	34,870	51	45	46	48	48%	20/96 (21%)
11	32,036	50	51	44	48	48%	50/96 (52%)
12	18,881	48	44	40	43	N/A	N/A

Reporting method: mean national individual percentiles.

Note: national averages are not 50. Norms vary a few percentiles on either side of 50, depending on subject and grade level.

Source: Statewide Report for Arizona Pupil Achievement Testing: June 1989

TEST SECURITY IN ARIZONA

Arizona changed test forms this year and administered the 1988 ITBS this spring at all grade levels. Arizona forbids delivery of tests to teachers earlier than the day of testing and they randomly monitor their assessments.

However, Arizona allows teachers to familiarize themselves with test questions (because of the publisher's recommendations), does not seal their test booklets, does not require special education students to be tested, does not use proctors, does not rotate questions yearly, and does not routinely scan answer sheets for suspicious erasures or analyze them for cluster variance.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

ARKANSAS* *SPRING 1989*

THE METROPOLITAN
ACHIEVEMENT TEST
FORM M 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
4	30,886	62	66	70	67	N/A	297/329(90%)
7	29,799	55	62	59	60	N/A	266/329(81%)
10	27,516	51	61	53	56	N/A	213/329(65%)

Reporting method: national percentile rank of mean normal curve equivalents
Source: Arkansas Standardized Testing Program Report

TEST SECURITY IN ARKANSAS

This is the fourth year that Arkansas has administered the Metropolitan Achievement Test and the state plans to administer it again this coming year. They also administer a criterion referenced test, the Minimum Performance Test.

Regulations required that districts deliver the tests to the schools no more than three days before testing and that principals give teachers the test booklets no earlier than noon the day before testing. In addition, random monitoring of test administration takes place.

However, teachers were not clearly forbidden from looking at the Metropolitan Achievement Test, the booklets were not sealed, proctors were not used, special education students were generally excluded from testing unless their IEP specifically recommended testing, and answer sheets were not scanned for suspicious erasures or analyzed for cluster variance.

Arkansas is revising its test security policies. They plan to prohibit teachers from looking at the test questions and to begin administering their norm-referenced tests in the fall. In addition, they hope to have the resources to seal their test booklets, to rotate questions every two years, and to scan answer sheets for suspicious erasures.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

CALIFORNIA* *MAY 1988*

THE CALIFORNIA ASSESSMENT PROGRAM EQUATED IN 1985 TO THE STANFORD THE CALIFORNIA AND THE CTBS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	328,013	282	284	281	N/A	N/A	N/A
6	301,334	265	273	270	N/A	N/A	N/A
8	285,116	252	263	264	N/A	N/A	N/A
12	226,269	250	N/A	250	N/A	N/A	N/A

Reporting method: scaled scores (all elementary grades are above the CTBS and Stanford national norms according to the equating studies; however, California no longer reports equivalent national percentiles)

Source: California Assessment Program Annual Report: 1987-88

TEST SECURITY IN CALIFORNIA

California Assessment Program security policies clearly forbid teachers to look at the test except as needed for administration, give teachers the test on the day of testing, randomly monitors their assessment and tests special education students unless their IEP prohibits it. In addition, California uses outside test proctors, administers multiple matrix forms of the test each year, and scans their answer sheets for suspicious erasures.

However, the booklets are not sealed and fall testing is not required. California is developing a new test for 1991 and they are considering expanding their program to give individual students' scores on a customized test.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

COLORADO* *FALL 1988* THE IOWA TEST OF BASIC SKILLS FORM G 1985 NATIONAL NORMS

GRADE	NUMBER TESTED*	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
4	11,569	62(58)	57(51)	60(53)	N/A	N/A	N/A
7	11,001	59(54)	55(49)	58(52)	N/A	N/A	N/A
10	8,890	60(58)	56(52)	58(57)	N/A	N/A	N/A

Reporting method: national individual percentile ranks. 1988 Form G national norm scores in parenthesis.

Source: Results of the 1988 Colorado Student Testing Program

* sample of students

TEST SECURITY IN COLORADO

Colorado administers The Iowa Test of Basic Skills in the fall to a representative sample of students to obtain state achievement data, not building or district data. Teachers obtain the test booklets only on the day of testing, testing is monitored by state officials, and questions are, in effect, rotated because a different sample of schools is tested every third year. Special education students are included in testing if they receive 50% or more of their instruction in regular classrooms.

However, teachers may look at test booklets, the booklets are not sealed, outside test proctors are not used, and answer sheets are not routinely scanned for suspicious erasures or routinely analyzed for cluster variance.

The Colorado State Board of Education recently adopted some new state regulations concerning the reporting methods that must be used by local school districts when reporting the results of their achievement tests. But, there are no state regulations that govern test security and test administration for the testing done in the local school districts.

CONNECTICUT *OCTOBER 1987*

THE CONNECTICUT MASTERY TEST EQUATED WITH METROPOLITAN ACHIEVEMENT TEST FORM ? 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
4	30,448	60	69	67	N/A	N/A	/162
6	28,954	57	65	66	N/A	N/A	/162
8	30,097	57	67	67	N/A	N/A	/162

Reporting method: mean individual national percentiles

Source: Connecticut State Department of Education.

TEST SECURITY IN CONNECTICUT

The Connecticut Mastery Test is administered in early fall. Test questions are changed every two years, special education students are generally included unless their IEP forbids testing, and the assessment is monitored by state officials.

Teachers are allowed to look at the test booklets, teachers may obtain the test booklets before the day of testing, the booklets are not sealed, outside test proctors are not routinely used, and answer sheets are not routinely scanned for suspicious erasures or routinely analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the testing done in the local school districts.

DELAWARE* *APRIL 1989*

THE STANFORD
ACHIEVEMENT TEST
FORM J 1988 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
1	9,944	48.4	50.3	50.7	49.7	47%	7/15(47%)
2	8,035	52.4	55.3	52.9	53.4	55%	13/15(87%)
3	7,595	52.8	54.6	52.0	53.4	54%	14/15(93%)
4	7,400	50.6	53.3	50.8	51.9	50%	12/15(80%)
5	6,877	51.2	52.2	49.7	51.4	51%	9/15(60%)
6	6,977	51.2	51.8	48.7	50.4	49%	7/15(47%)
7	6,866	50.7	50.3	49.7	50.2	49%	10/16(62%)
8	6,671	51.0	51.5	50.5	50.7	50%	9/16(56%)
11	5,657	51.6	51.7	51.3	51.6	51%	11/17(64%)

Reporting method: mean normal curve equivalents

Source: Delaware Educational Assessment Program Statewide Test Results Summary Report: Spring 1989

TEST SECURITY IN DELAWARE

Delaware administered the Stanford Achievement Test, Form J, for the first time in April of 1989. The contract with the publisher is for a two year period. Delaware policy states teachers may obtain the test booklets no earlier than 24 hours before the day of testing, the testing is routinely monitored by state officials, and special education students are generally tested unless their IEP prohibits testing.

Teachers are allowed to look at the test booklets, the booklets are not sealed and fall testing is not required. In addition, outside test proctors are not routinely used, and answer sheets are not routinely scanned for suspicious erasures or analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the testing done independently in the local school districts.

FLORIDA* *SPRING 1988*

FLORIDA NORM-REFERENCED
TESTING PROGRAM
SUBSET OF QUESTIONS FROM NAEP
1982 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	113,050	41.47	N/A	N/A	N/A	N/A	45/67(67%)
7	104,857	50.97	N/A	N/A	N/A	N/A	37/67(55%)

Reporting method: mean student scaled scores
National 3rd grade mean score equals 40.81
National 7th grade mean score equals 50.61
Source: Florida State Department of Education

TEST SECURITY IN FLORIDA

Florida administers a graduation exam and criterion referenced test: the State Student Assessment Test. Their Norm-Referenced Testing Program was a subset of NAEP items, but it will no longer be used because of federal regulations.

In Florida, teachers are clearly prohibited from examining test booklets, teachers may not obtain the test booklets before the day of testing, test booklets are sealed either by shrink wrapping them in class size packets, or by individually sealing the actual booklets. Testing is routinely monitored by state officials, but visits are announced. In addition, special education students are generally tested unless they participate 12 hours or less in regular classrooms. One-third of questions are new every year, and answer sheets are scanned for suspicious erasures.

However, special education student's scores are excluded from district's results, outside test proctors are not routinely used and answer sheets are not analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the testing done independently in the local school districts.

GEORGIA*

MARCH 1989 THE IOWA TEST OF BASIC SKILLS
FORM G 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
2	87,062	63	73	74	70	67.6%	173/186(93.0%)
4	81,320	54	62	64	59	59.0%	133/186(71.9%)
7	73,415	51	55	55	55	55.6%	105/185(56.7%)

TEST OF ACHIEVEMENT AND PROFICIENCY
FORM G 1985 NATIONAL NORMS

9	72,661	52	57	50	54	55.6%	115/178(64.6%)

Reporting method: mean individual percentile scores
Source: Georgia Student Assessment Program Official State Summary 1988-89

TEST SECURITY IN GEORGIA

Georgia has administered the Iowa Test of Basic Skills, Form G, for the last three years. They will administer a new test, Form J of the ITBS, during the 1989-90 school year. They also administer a criterion referenced exam and a graduation exam.

The Iowa Test of Basic Skills booklets are given to teachers on the day of testing, teachers are not allowed to look at the Iowa test booklets except as needed during administration, the testing is routinely monitored by state officials, special education students are tested unless their IEP specifically prohibits testing, and proctors are used.

However, schools may obtain the test booklets before the day of testing, the booklets are not sealed, test questions have not been rotated every year, and answer sheets are not routinely scanned for suspicious erasures or analyzed for cluster variance.

There are no state regulations that govern test security for the testing done independently in the local Georgia school districts.

HAWAII *SPRING 1988* THE STANFORD ACHIEVEMENT TEST FORM E 1982 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	12,971	77(42)	N/A	81(77)	(N/A)	N/A	3/7(43%)
6	12,150	80(55)	N/A	81(82)	(N/A)	N/A	6/7(86%)
8	11,023	75(52)	N/A	73(59)	(N/A)	N/A	5/7(71%)
10	9,611	76(46)	N/A	77(71)	(N/A)	N/A	2/7(29%)

Reporting method: Percentage of children in stanine 4 or above (national group percentile ranks in parenthesis)

Source: Administrator of Testing—Hawaii Department of Education

TEST SECURITY IN HAWAII

Hawaii administers the Stanford Achievement Test as well as a criterion referenced test: the Hawaii State Test of Essential Competencies. Only the student at the time of testing may look at the criterion referenced test. Hawaii is in the process of implementing stricter security measures, including random monitoring of testing by state officials.

However, teachers are allowed to look at the Stanford Achievement Test booklets. In addition, teachers may be given the test booklets well before the day of testing and the booklets are not sealed. Special education students are excluded unless their IEP allows testing, proctors are not used, test questions have not been rotated, and answer sheets are not routinely scanned for suspicious erasures or analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the testing done independently in the local Hawaiian school districts.

IDAHO* *FEBRUARY 1988* THE IOWA TEST OF BASIC SKILLS FORM G 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
6	15,073	72	62	75	69	67%	N/A
8	14,193	68	70	71	70	67%	N/A
11	12,854	80	77	67	75	68%	N/A

Reporting method: group (school) national percentile ranks

Source: 1988 Performance Summary Idaho Statewide Testing Program.

Note: Idaho reports with group percentile ranks when issuing press releases.

TEST SECURITY IN IDAHO

This is the fourth year that Idaho has administered the Iowa Test of Basic Skills, Form G, to its eleventh graders, the third year it has been given to its eighth graders and the first year it was given to its sixth graders. There are no plans to change tests. Teachers are handed the test on the day of testing and special education students are included in testing unless their IEP specifically forbids testing.

However, teachers are allowed to look at the test booklets, the booklets are not sealed, fall testing is not required, the testing is not routinely monitored by state officials, and outside test proctors are not routinely used. In addition, test questions have not been rotated every year, answer sheets are not routinely scanned for suspicious erasures or analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the testing done independently in the local school districts.

ILLINOIS* *APRIL 1988* ILLINOIS GOAL ASSESSMENT PROGRAM EQUATED WITH STANFORD ACHIEVEMENT TEST FORM F 1986 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50*	% DISTRICTS > OR = 50
3	121,330	59	N/A	N/A	N/A	59%	803/862 (93%)
6	108,799	55	N/A	N/A	N/A	55%	737/854 (86%)
8	107,791	50	N/A	N/A	N/A	50%	575/850 (68%)

Reporting method: percentage of students scoring above national norm

Source: Illinois State Board of Education

* In reading

TEST SECURITY IN ILLINOIS

The Illinois Goal Assessment Program (IGAP) is normed by the concurrent administration of the Stanford Achievement Test to a sample of students. The 1989 test was normed with the mathematics subtest of the Stanford Achievement test, Form K. Most of the IGAP questions are new every year and the tests are shrunk-wrapped for delivery to the schools.

However, teachers are allowed to look at IGAP test booklets, teachers may obtain the test booklets before the day of testing, fall testing is not required, and the testing is not routinely monitored by state officials. In addition, special education students are generally excluded unless their IEP recommends testing, outside test proctors are not routinely used, and answer sheets are not routinely scanned for suspicious erasures or analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the testing done independently in the local school districts.

INDIANA *MARCH 1988* THE CALIFORNIA ACHIEVEMENT TEST FORM E 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
1	75,549	65	68	78	N/A	62.9%	300/301 (99%)
2	71,361	64	68	73	70	71.1%	299/301 (99%)
3	70,316	64	71	72	70	70.3%	297/301 (99%)
6	65,861	63	64	69	66	65.6%	289/300 (96%)
8	65,549	60	61	67	63	62.4%	283/300 (94%)
9	65,345	57	58	62	60	57.8%	261/294 (89%)
11	62,321	58	59	63	60	59.4%	254/294 (86%)

Reporting method: median individual percentile ranks

Source: Indiana Statewide Testing For Educational Progress Report
State of Indiana, Department of Education

TEST SECURITY IN INDIANA

This is the second year the California Achievement Test has been administered in Indiana. The state monitors the testing and special education students are tested unless their IEP forbids it.

However, teachers are allowed to look at the test booklets, teachers may obtain the test booklets before the day of testing, the booklets are not sealed, fall testing is not required, and outside test proctors are not routinely used. In addition, test questions have not been rotated every year, and answer sheets are not routinely scanned for suspicious erasures or routinely analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the testing done independently in the local school districts.

IOWA *1988-89*

THE IOWA TEST OF BASIC SKILLS
FORM G&H 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50*
K	N/A	N/A	N/A	N/A	68	68%	374/435 (86%)
1	N/A	N/A	N/A	N/A	74	74%	391/435 (90%)
2	N/A	N/A	N/A	N/A	72	72%	396/435 (91%)
3	N/A	N/A	N/A	N/A	71	71%	391/435 (90%)
4	N/A	N/A	N/A	N/A	72	72%	400/435 (92%)
5	N/A	N/A	N/A	N/A	73	73%	409/435 (94%)
6	N/A	N/A	N/A	N/A	73	73%	418/435 (96%)
7	N/A	N/A	N/A	N/A	73	73%	418/435 (96%)
8	N/A	N/A	N/A	N/A	70	70%	422/435 (97%)

Reporting method: percentage above national median

Source: Iowa Basic Skills Testing Program Score Reports and Norms 1988-89

* estimated from the percentage of schools above the national norm

TEST SECURITY IN IOWA

Testing in Iowa is under the control of the local districts; however, more than 95% of Iowa school districts administer the ITBS. The University of Iowa oversees administration of the ITBS in Iowa and rotates test forms every year for those districts participating. In addition, most testing is done in the fall.

There are no state regulations that govern test security and test administration for the nationally normed testing done independently in the local school districts. No state regulations prevent teachers from looking at test booklets, teachers may obtain the test booklets before the day of testing, and the booklets are not required to be sealed. In addition, fall testing is not required, the testing is not routinely monitored by state officials, special education students are not specifically included, outside test proctors are not required, and answer sheets are not routinely scanned for suspicious erasures or analyzed for cluster variance.

KANSAS* *SPRING 1989*

VARIOUS NATIONALLY NORMED TESTS

DISTRICT	TEST	ELEMENTARY	SECONDARY
Wichita	ITBS-G	"Above 50"	"Above 50"
Shawnee-Mission	ITBS-G	"Above 50"	"Above 50"
Kansas City	CTBS-U/V	"Above 50"	"Above 50"
Topeka	ITBS-7	"Above 50"	"Above 50"
Olathe	ITBS-G	"Above 50"	"Above 50"
Lawrence	SRA	"Above 50"	"Above 50"
Salina	SRA	"Above 50"	"Above 50"
Blue Valley	ITBS-G	"Above 50"	"Above 50"
Junction City	CAT-E	"Above 50"	"Above 50"
Manhattan	ITBS-G	"Above 50"	"Above 50"

Source: telephone survey, Spring 1989

TEST SECURITY IN KANSAS

Kansas has administered the Kansas Minimum Competency Test for the last ten years, but it is not nationally normed. Some of the test questions have been rotated every year, special education students are tested unless they are enrolled only in non-academic classes, and test booklets have been sealed. However, it will not be administered in 1989-1990.

Kansas does not administer any statewide nationally normed tests. However, norm-referenced tests are used independently in the districts and there are no state regulations that govern test security and test administration for the norm-referenced testing done in the local districts. No state regulations prevent teachers from looking at norm-referenced test booklets, teachers may obtain the test booklets before the day of testing, the booklets are not required to be sealed, and fall testing is not required. In addition, norm-referenced testing is not routinely monitored by state officials, special education students are not specifically included, outside test proctors are not required, test questions do not have to be rotated every year, and answer sheets are not necessarily scanned for suspicious erasures or analyzed for cluster variance.

Kansas is one of the eleven states that is not planning to participate in the 1990 National Assessment of Educational Progress state-by-state evaluation.

KENTUCKY* *APRIL 1988* KENTUCKY ESSENTIAL SKILLS TEST (KEST) MODIFICATION OF CTBS FORM U 1981 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
1	52,609	55.1	57.3	67.7	N/A	N/A	N/A
2	48,828	55.9	59.1	73.2	62.6	72.4%	178/178 (100%)
3	46,416	60.8	59.7	60.2	62.5	74.7%	177/178 (99%)
4	45,032	59.0	53.4	60.4	57.1	62.0%	166/178 (93%)
5	44,882	53.7	57.5	60.8	56.2	60.1%	172/178 (97%)
6	42,163	57.3	58.5	64.0	60.3	69.1%	177/178 (99%)
7	42,430	56.2	60.1	59.8	58.5	66.7%	174/178 (98%)
8	39,620	55.1	59.1	60.6	57.4	64.1%	171/178 (96%)
9	43,085	46.7	53.9	54.3	51.7	52.5%	122/175 (67%)
10	41,939	48.3	57.0	55.8	54.5	59.0%	145/174 (83%)
11	40,830	48.0	55.1	51.9	52.0	53.8%	115/173 (66%)
12	36,207	45.5	54.2	51.0	51.1	53.3%	96/173 (55%)

Reporting method: normal curve equivalents

Source: Kentucky Essential Skills Test, Statewide Testing Results

TEST SECURITY IN KENTUCKY

Kentucky has administered the KEST for the last four years. They will administer the new CTBS test during the 1989-1990 school year. Special education students are tested unless their IEP specifically forbids testing. A new test security policy has recently been adopted but it does not cover the most common forms of testing irregularities.

Teachers are allowed to look at the test booklets under controlled conditions, teachers may obtain the test booklets before the day of testing, the booklets are not sealed, fall testing is not required and the testing is not routinely monitored by state officials. In addition, outside test proctors are not routinely used, test questions have not been rotated every year, and answer sheets are not routinely scanned for suspicious erasures or routinely analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the testing done independently in the local school districts.

LOUISIANA* *APRIL 1989*

THE CALIFORNIA
ACHIEVEMENT TEST
FORM F 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
4	57,364	47.1	43.2	48.2	41.6	39.5	12/67(18%)
6	52,607	46.1	45.0	51.2	46.9	45.0	22/67(33%)
9	45,068	42.8	45.6	47.2	41.3	41.3	8/67(12%)

Reporting method: median individual national percentiles
Source: Louisiana Statewide Norm-Referenced Test Results: Spring 1989

TEST SECURITY IN LOUISIANA

Louisiana also administers a criterion referenced test and a graduation exam. This was the third year of statewide testing with the California Achievement Test, Form F. Special education students in Louisiana are tested unless their IEP specifically forbids testing.

However, teachers are allowed to look at norm-referenced test booklets, teachers may obtain the test booklets before the day of testing, the booklets are not sealed, and fall testing is not required. In addition, the testing is not routinely monitored by state officials, outside test proctors are not routinely used, test questions have not been rotated every year, answer sheets are not routinely scanned for suspicious erasures or analyzed for cluster variance.

Our 1987 survey revealed that the majority of districts in Louisiana use additional nationally normed tests and that most were testing above the national norm. There are no state regulations that govern test security and test administration for the testing done independently in the local school districts.

MAINE* *1987-88*

MAINE EDUCATIONAL ASSESSMENT

GRADE	NUMBER TESTED	READING	LANG	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
4	14,012	250(+13.5)	N/A	255(+6.9)	N/A	N/A	N/A/158
8	14,723	255(+3.7)	N/A	285(-0.9)	N/A	N/A	N/A/158
11	15,020	265(+6.2)	N/A	255(+1.3)	N/A	N/A	N/A/158

Reporting method: scaled scores (differences from nation in average percent correct for 1986 scores in parenthesis.
All scores are higher now.)
Source: Maine Educational Assessment 1987-88 State Summary and Interpretations Report

TEST SECURITY IN MAINE

The Maine Educational Assessment is reported with scaled scores and is no longer equated with nationally normed tests. Teachers may not obtain the test booklets until the day before testing, the booklets are shrink-wrapped when delivered to the schools, and testing is monitored by state officials. The eighth grade is tested in the fall. In addition, special education students are generally included in testing, outside test proctors are used, test questions are rotated via matrix sampling, answer sheets are routinely scanned for suspicious erasures and analyzed for cluster variance.

However, there are no state regulations that govern test security and test administration for the norm-referenced testing conducted independently in the local school districts.

Maine is one of the eleven states that is not planning to participate in the 1990 National Assessment of Educational Progress state-by-state evaluation.

MARYLAND* *FALL 1988*

THE CALIFORNIA ACHIEVEMENT TEST
FORM C 1978 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	52,000	3.7	3.9	3.5	N/A	N/A	23/24(96%)
5	47,000	6.1	7.3	6.0	N/A	N/A	24/24(100%)
8	44,000	10.0	10.3	9.9	N/A	N/A	24/24(100%)

Reporting method: grade equivalent scales (all of the above scores are well above the publisher's national norms).

Source: Maryland Accountability Testing Program Annual Report 1988-89

TEST SECURITY IN MARYLAND

Maryland has used the 1978 California Achievement Test, Form C, since 1980 and they plan to administer it again in the fall of 1989. Maryland will adopt a new test for administration in the spring of 1991. Maryland also administers graduation exams in reading, mathematics, citizenship, and writing.

Maryland has a "Code of Ethics" that prohibit either the graduation exam or the norm-referenced test from being given to teachers earlier than the day of testing. In addition, special education students are generally tested under new regulations unless their IEP prohibits testing, and teachers are prohibited from looking at test booklets except as needed during administration. The norm-referenced test is administered in the fall. The graduation exam rotates most questions every year. Maryland is now revising its test exclusion and test security policies.

If adequate funding is available, Maryland plans to seal norm-referenced test booklets, and begin random, unannounced, monitoring of its administration. They also hope to be able to rotate test questions by using equivalent forms, scan answer sheets for suspicious erasures, and analyze them for cluster variance.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

MASSACHUSETTS*

APRIL 1986 — MASSACHUSETTS EDUCATIONAL ASSESSMENT PROGRAM SUBSET OF NAEP ITEMS 1986 NATIONAL NORMS (NAEP)

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	51,949	1300	N/A	1300	N/A	63%	219/247(89%)
7	55,168	1300	N/A	1300	N/A	62%	210/237(89%)
11	60,217	1300	N/A	1300	N/A	57%	172/255(67%)

Reporting method: scaled scores (all math scores are above the national NAEP averages, national reading averages are not available)

Source: The Massachusetts Educational Assessment Program 1988 Summary

TEST SECURITY IN MASSACHUSETTS

Massachusetts administers a basic skills test, the Massachusetts Basic Skills Testing Program and the Massachusetts Educational Assessment Program. The former test is administered to all students in the 3rd, 6th and 9th grades while the latter test uses a matrix sampling technique to obtain school and district data. Both tests are delivered to schools sealed in a shrink wrapped package, special education students are generally tested, questions are rotated every year, and answer sheets are routinely scanned for suspicious erasures. Massachusetts is now revising their test security and is creating a Code of Testing Ethics.

As of now, teachers are allowed to look at the test booklets, teachers may obtain the test booklets before the day of testing, fall testing is not required for all tests, the testing is not routinely monitored by state officials, outside test proctors are not routinely used, and answer sheets are not routinely analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

MICHIGAN* *SPRING 1989*

VARIOUS NATIONALLY NORMED TESTS

DISTRICT	TEST	ELEMENTARY	SECONDARY
Detroit	CAT-E	"Below 50"	"Below 50"
Flint	ITBS-G	"Below 50"	"Below 50"
Grand Rapids	CAT-E	"Above 50"	"Below 50"
Lansing	SAT-7	"Above 50"	"Above 50"
Utica	ITBS	"Above 50"	"Above 50"
Wayne Westland	CAT-E	"Above 50"	"unavailable"
Pontiac	CAT-E	"Below 50"	"Below 50"
Warren	ITBS-7	"Above 50"	"Above 50"
Saginaw	CAT-E	"Above 50"	"Below 50"
Livonia	CTBS-U	"Above 50"	"Above 50"

Source: Telephone Survey, Spring 1989

TEST SECURITY IN MICHIGAN

The state administers the Michigan Educational Assessment Program, a criterion referenced test, to grades 4, 7 and 10. It is not nationally normed; scores are reported as percent mastery. The test is administered in the early fall, the booklets are shrunk-wrapped in class size packets, the testing is routinely monitored by state officials, special education students are generally tested, and answer sheets are scanned for suspicious erasures and analyzed for cluster variance.

However, teachers are allowed to look at the test booklets, and teachers may obtain the test booklets before the day of testing. In addition, outside test proctors are not routinely used, and test questions have not been rotated every year. However, questions will be changed on a yearly basis on the new Michigan Educational Assessment Program which will first be administered in the fall of 1989.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

MINNESOTA* *Spring 1989*

VARIOUS NATIONALLY NORMED TESTS

DISTRICT	TEST	ELEMENTARY	SECONDARY
Minneapolis	SAT&CAT	"Above 50"	"Above 50"
Anoka	ITBS	"Above 50"	"Above 50"
St. Paul	SRA	"Above 50"	"Above 50"
Osseo	CTBS-U	"Above 50"	"Above 50"
Rosemount	MAT-6	"Above 50"	"Above 50"
Bloomington	SAT	"Above 50"	"Above 50"
Moundsview	CTBS-U/V	"Above 50"	"Above 50"
St. Cloud	CTBS-U	"Above 50"	"Above 50"
Rochester	CAT-E	"Above 50"	"Above 50"
Stillwater	CAT-E	"Above 50"	"Above 50"

Source: Telephone Survey, Spring 1989

TEST SECURITY IN MINNESOTA

The state administers the Minnesota Assessment Test, a criterion referenced test, at various grades. The test is not nationally normed. Tests are rotated every year and outside test proctors are used. Other details are not available.

Minnesota does not administer any statewide, nationally normed test, but they are used extensively in the local districts. There are no state regulations that govern test security for the testing done independently in the local school districts. No state regulations prevent teachers from looking at norm-referenced test booklets, teachers may obtain the test booklets before the day of testing, the booklets are not required to be sealed, and fall testing is not required. In addition, the testing is not routinely monitored by state officials, special education students are not specifically included, outside test proctors are not required, test questions do not have to be rotated every year, and answer sheets are not necessarily scanned for suspicious erasures or analyzed for cluster variance.

MISSISSIPPI* *APRIL 1988*

THE STANFORD
ACHIEVEMENT TEST
FORM E 1982 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	BASIC BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
1	44,634	55.1	51.9	51.9	53.7	N/A	116/152(76%)
4	37,430	48.4	52.3	51.1	49.9	N/A	71/152(47%)
6	34,661	48.7	52.6	51.7	50.2	N/A	66/152(43%)

Reporting method: mean national normal curve equivalents
Source: Mississippi Statewide Testing Pupil Performance 1988

TEST SECURITY IN MISSISSIPPI

Mississippi administers three statewide tests: a norm-referenced test (the Stanford), a criterion referenced test, and a graduation exam. The last two tests are given to teachers on the day of testing, both are sealed, 80% of the questions are new every year, and the administration of the criterion referenced test and the graduation exam are routinely monitored by state officials.

Regulations require that the Stanford Achievement Test be delivered to teachers on the day of testing and most special education students will be tested. They plan to clearly tell teachers not to read the Stanford, the testing will be routinely monitored by state officials, and they plan to scan answer sheets for suspicious erasures and analyze them for cluster variance. Mississippi will administer the new, secure version of the Stanford Achievement Test in 1989-1990.

However, the booklets will not be sealed, fall testing will not be required, test questions will not be rotated every year, and outside test proctors will not be routinely used.

There are no state regulations that govern test security for the norm-referenced testing done independently in the local school districts.

Mississippi is one of the eleven states that is not planning to participate in the 1990 National Assessment of Educational Progress state-by-state evaluation.

MISSOURI* *SPRING 1989* THE IOWA TEST OF BASIC SKILLS FORM G/H 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	5706	53	64	59	N/A	N/A	N/A
6	5152	54	59	62	N/A	N/A	N/A
8	4733	50	57	59	N/A	N/A	N/A
10	5023	51	N/A	52	N/A	N/A	N/A

Reporting method: median national individual percentiles of the "state sample."
Source: Summary of 1988 Results from the Missouri Mastery and Achievement Test

TEST SECURITY IN MISSOURI

The Missouri Mastery and Achievement Test is a criterion referenced test that is administered statewide to grades 2 through 10. The Iowa Test of Basic Skills is administered concurrently to a stratified random sample of students every year. The tests are sealed, some questions are new every year, and special education students are generally tested. New regulations being considered in Missouri would prohibit teachers from looking at test booklets except as needed during administration, and prohibit teachers from receiving the test booklets earlier than the day of testing.

However, fall testing is not required, the testing is not routinely monitored by state officials, outside test proctors are not routinely used, and answer sheets are not routinely scanned for suspicious erasures or routinely analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

Missouri is one of the eleven states that is not planning to participate in the 1990 National Assessment of Educational Progress state-by-state evaluation.

MONTANA* *SPRING 1989*

VARIOUS NATIONALLY NORMED TESTS

DISTRICT	TEST	ELEMENTARY	SECONDARY
Billings	ITBS-G	"Above 50"	"Above 50"
Great Falls	ITBS-G	"Above 50"	"Above 50"
Helena	CTBS-U	"Above 50"	"Above 50"
Butte	SAT-E	"Above 50"	"Above 50"
Missoula	SRA	"Above 50"	"Average"
Kalispell	ITBS-G	"Above 50"	"Above 50"
Columbia Falls	ITBS-G	"Above 50"	"Above 50"
Bozeman	SAT-E	"Above 50"	"Above 50"

Source: Telephone Survey, Spring 1989

TEST SECURITY IN MONTANA

Montana does not currently administer any statewide test. They are in the process of implementing a statewide norm-referenced testing program by requiring all 600+ districts in Montana to use one of six norm-referenced tests. The law mandates that the testing be done in the spring and "special education students shall not be required to participate."

At the present time, there are no state regulations that govern test security and test administration for the testing that will be required in the local school districts. No state regulations prevent teachers from looking at norm-referenced test booklets, teachers may obtain the test booklets before the day of testing, and the booklets are not required to be sealed. Fall testing will not be required, the testing will not be routinely monitored by state officials, special education students will not specifically be included, outside test proctors will not be required, test questions do not have to be rotated every year, and answer sheets will not necessarily be scanned for suspicious erasures or analyzed for cluster variance.

NEBRASKA* *SPRING 1989*

VARIOUS NATIONALLY NORMED TESTS

DISTRICT	TEST	ELEMENTARY	SECONDARY
Omaha	CAT-E	"Above 50"	"Below 50"
Lincoln	CAT-E	"Above 50"	"Above 50"
Millard	CAT-E	"Above 50"	"Above 50"
Bellevue	ITBS-G	"Above 50"	"Above 50"
Grand Island	CTBS-U	"Above 50"	"Above 50"
Papillion/La Vista	CAT-E/F	"Above 50"	"Above 50"
Westside	CTBS-U	"Above 50"	"Above 50"
South Sioux City	SRA	"Above 50"	"Above 50"
North Platte	SRA	"Above 50"	"Above 50"
Scottsbluff	SRA	"Above 50"	"Above 50"

Source: Telephone Survey, Spring 1989

TEST SECURITY IN NEBRASKA

Nebraska has no statewide testing program but they plan to start collecting local district results in the 1989-1990 school year.

There are no state regulations that govern test security and test administration for the tests that will be compiled by the state. No state regulations will prevent teachers from looking at norm-referenced test booklets, teachers may obtain the test booklets before the day of testing, the booklets are not required to be sealed, and fall testing is not required. The norm-referenced testing will not be routinely monitored by state officials, special education students will not specifically be included, outside test proctors are not required, test questions do not have to be rotated every year, and answer sheets will not necessarily be scanned for suspicious erasures or routinely analyzed for cluster variance.

NEVADA* 1987-88

THE STANFORD ACHIEVEMENT TEST
FORM E 1982 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50*	% DISTRICTS > OR = 50*
3	11,902	91(65)	94(62)	92(65)	N/A	N/A	16/17(94%)
6	10,521	85(56)	84(59)	87(62)	N/A	N/A	16/17(94%)

SCIENCE RESEARCH ASSOCIATES
FORM ? 1985 NATIONAL NORMS

9	10,006	59	N/A	63	N/A	N/A	N/A

Reporting method: Percentage of students in average and above average stanines for grades 3 and 6 (estimated national percentile ranks in parenthesis). Grade 9 is reported as the percentage of students above the 50th percentile.

* in language from 1987 district reports in Nevada Education in 1988, A Status Report.

Source: The Nevada Proficiency Examination Program Results of the 1987/88 Examinations

TEST SECURITY IN NEVADA

Nevada has administered the Stanford Achievement Test, Form E, for the last four years. Nevada also administers state-developed proficiency examinations in reading, mathematics, and writing, which are required for graduation.

New legislation will require all school districts to administer the reading and math subtests of the CTBS-4 at grades 3, 6 and 9 during the 1990-91 school year. Any student performing at or below the 22nd percentile will be considered to have "not demonstrated adequate achievement." Scores will be reported at the building and district levels in national percentiles, normal curve equivalents, and stanines.

A Nevada statute prohibits the disclosure of test content, but it does not prohibit teachers from obtaining and studying test booklets in advance of testing. Special education children are tested if they have been mainstreamed into regular classrooms, but other special education students may be excluded from testing.

The Stanford Achievement Test booklets are not sealed and fall testing is not specifically mandated. In addition, the testing is not routinely monitored by state officials, outside test proctors are not required, test questions do not have to be rotated every year, and answer sheets are not scanned for suspicious erasures or routinely analyzed for cluster variance.

NEW HAMPSHIRE* *SEPTEMBER 1988* THE CALIFORNIA ACHIEVEMENT TEST FORM E 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
4	12,457	63.7	59.1	62.3	61.9	N/A	120/147 (82%)
8	11,366	62.3	57.6	61.7	60.1	N/A	101/121 (83%)
10	10,328	56.4	52.6	61.5	61.5	N/A	53/74 (72%)

Reporting Method: median national individual percentiles

Source: New Hampshire Statewide Testing Program 1989 Summary Report

TEST SECURITY IN NEW HAMPSHIRE

The California Achievement Test, Form E, has been administered in the early fall for the last four years, and they will administer it again next year. New Hampshire did not supply us with copies of their written test security procedures, but they assert that tests are delivered to teachers shortly before testing, districts are asked not to allow teachers to look at the test booklets prior to administration, and that the tests are sealed.

However, the testing is not routinely monitored by state officials, special education students are generally excluded unless they are mainstreamed into regular classrooms 50 percent or more of the time, outside test proctors are not routinely used, and test questions have not been rotated every year. In addition, answer sheets are not routinely scanned for suspicious erasures or analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

NEW JERSEY *SPRING 1989*

VARIOUS NATIONALLY NORMED TESTS

DISTRICT	TEST	ELEMENTARY	SECONDARY
Newark	CTBS-U	"Below 50"	"Below 50"
Jersey City	MAT-6	"Below 50"	"Below 50"
Patterson	CAT-E	"Above 50"	
Camden	CTBS-U	"Above 50"	"Above 50"
Toms River	ITBS-G/H	"Above 50"	"Average"
Elizabeth	CAT-E	"Above 50"	Refused
Trenton	CAT-E	"Above 50"	"Below 50"
East Orange	CAT-E	"Above 50"	"Below 50"
Hamilton	CAT-E/F	"Above 50"	"Above 50"
Cherry Hill	ITBS-G	"Above 50"	"Above 50"

Source: Telephone Survey, Spring 1989

TEST SECURITY IN NEW JERSEY

New Jersey does not administer a statewide norm-referenced test, but they do require that all 600+ school districts administer a norm-referenced test to identify children needing remediation. The state sets minimum performance standards for these tests.

New Jersey administers a statewide graduation exam with extensive test security procedures. The graduation test booklet is sealed, tests are given to teachers on the day of testing, questions are new every year, teachers are told not to look at the test booklets, and special education students are included in the graduation testing.

However, there are no state regulations that govern test security and test administration for the nationally normed testing that the state requires in the local school districts. No state regulations prevent teachers from looking at norm-referenced test booklets and teachers may obtain the test booklets before the day of testing. In addition, the booklets are not required to be sealed, the testing is not routinely monitored by state officials, outside test proctors are not required, test questions do not have to be rotated every year, and answer sheets are not necessarily scanned for suspicious erasures or routinely analyzed for cluster variance

NEW MEXICO

MARCH 1988 THE COMPREHENSIVE TEST OF BASIC SKILLS
FORM U 1981 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	21,486	54	63	66	61	N/A	77/90 (86%)
5	20,250	57	60	63	59	N/A	73/90 (81%)
8	19,009	56	64	61	57	N/A	71/90 (79%)

Reporting method: national individual percentile ranks

Source: New Mexico Articulated Assessment System Student Achievement Profile: 1987-1988

TEST SECURITY IN NEW MEXICO

This was the fifth year New Mexico administered the CTBS-U. A new edition of the CTBS was administered this year, in which 50% of the items were locally developed and 50% were CTBS items. New Mexico also administers a criterion referenced test and a high school graduation exam.

The graduation exam is sealed, questions are changed every year, teachers are told not to look at the exam, and the graduation exam is delivered to schools shortly before testing.

On the norm-referenced test, teachers are allowed to look at the test booklets, teachers may obtain the test booklets before the day of testing, the booklets are not sealed, fall testing is not required, and testing is not routinely monitored by state officials. In addition, special education students are generally excluded unless their IEP recommends testing, outside test proctors are not routinely used, test questions have not been rotated every year, and answer sheets are not routinely scanned for suspicious erasures or analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

NEW YORK* *SPRING 1989*

VARIOUS NATIONALLY NORMED TESTS

DISTRICT	TEST	ELEMENTARY	SECONDARY
New York City	MAT-6(math)	"Above 50"	"Above 50"
	DRP(reading)	"Below 50"	"Below 50"
Buffalo	SAT-F	"Above 50"	"Above 50"
Rochester	CAT-E	"Above 50"	"Below 50"
Syracuse	ITBS	"Average"	"Average"
Yonkers	MAT-6	"Above 50"	"Above 50"
Sachem	SAT	"Above 50"	
Middle Country	SAT-86	"Average"	"Above 50"
Greece Cen.	CAT-E	"Above 50"	"Above 50"
E. Ramapo Cen.	ITBS	"Above 50"	"Above 50"

Source: Telephone Survey, Spring 1989

TEST SECURITY IN NEW YORK

New York does not administer any nationally normed tests. They do administer a number of state-developed criterion referenced tests: the Pupil Evaluation Program (5 tests administered in grades 3, 5, and 6), the Regents Competency Test (6 tests administered in grades 9-12), and the Regents Examinations (17 tests administered in grades 9-12). These tests are sealed, teachers are not allowed to look at the test booklets prior to administration, teachers may not obtain the test booklets before the day of testing, the testing is routinely monitored by state officials, each test is composed of questions that have never been used before, outside test proctors may be used, and tests are shipped in locked metal chests.

However, there are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

New York is one of the eleven states that is not planning to participate in the 1990 National Assessment of Educational Progress state-by-state evaluation.

NORTH CAROLINA *SPRING 1988*

THE CALIFORNIA ACHIEVEMENT TEST
FORM E 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	80,944	54	61	65	60	N/A	133/142(94%)
6	77,608	51	58	58	55	N/A	115/142(81%)
8	81,426	51	55	55	53	N/A	100/130(77%)

Reporting method: median national individual percentiles
Source: Report of Student Performance, 1986-1988 Division of Testing, North Carolina Department of Public Instruction.

TEST SECURITY IN NORTH CAROLINA

North Carolina has administered the California Achievement Test Form E for last four years. They also administer a high school competency exam.

North Carolina would not supply us with copies of their test procedure or test security policies. William J. Brown Jr., the director of North Carolina's "Division of Accountability Services," also refused to verify any data for North Carolina, although he claimed he found "glaring errors" in our data. We obtained the above scores directly from their own publication and rechecked it twice, but we are unable to supply details of their state test security policies without written copies of their test security policies.

We know of no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

NORTH DAKOTA

SEPTEMBER 1988 THE IOWA TEST OF BASIC SKILLS
FORM G 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	2,902	66	64	63	66	N/A	N/A
5	3,407	68	63	67	67	N/A	N/A
7	3,003	65	55	67	66	N/A	N/A
9	2,627	63	65	67	69	N/A	N/A
11	2,303	65	65	66	67	N/A	N/A

Reporting method: mean individual national percentiles

SCIENCE RESEARCH ASSOCIATES
FORM P 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	2,315	75	73	66	72	N/A	N/A
5	2,335	71	69	70	71	N/A	N/A
7	2,747	69	70	69	70	N/A	N/A
9	3,108	65	64	75	69	N/A	N/A
11	3,171	61	62	72	66	N/A	N/A

Reporting method: composite percentile scores
Source: North Dakota Fall Statewide Achievement Test Data

TEST SECURITY IN NORTH DAKOTA

North Dakota does not administer any statewide tests. However, they collect data from their districts, the majority of which use one of the above two tests. They plan to administer a statewide test in the 1989-1990 school year.

There are no state regulations that govern test security and test administration for the nationally normed testing done in the local school districts. No state regulations prevent teachers from looking at norm-referenced test booklets, teachers may obtain the test booklets before the day of testing, the booklets are not required to be sealed, and fall testing is not required. The testing done in the districts is not routinely monitored by state officials, special education students are not required to be tested, outside test proctors are not required, test questions do not have to be rotated every year, and answer sheets are not necessarily scanned for suspicious erasures or routinely analyzed for cluster variance.

OHIO* *SPRING 1989* VARIOUS NATIONALLY NORMED TESTS

DISTRICT	TEST	ELEMENTARY	SECONDARY
Cleveland	CTBS	"Below 50"	"Below 50"
Columbus	CTBS-V	"Above 50"	"Above 50"
Cincinnati	CAT-E/F	"Below 50"	"Below 50"
Toledo	ITBS	"Above 50"	"Above 50"
Akron	CAT-E	"Above 50"	"Below 50"
S. Westaria City	CTBS-U	"Above 50"	"Above 50"
Youngstown City	SAT-7	"Below 50"	"Below 50"
Canton	ITBS-G	"Above 50	"Below 50"
Parma City	CAT-E/F	"Above 50"	"Above 50"
Loraine City	CTBS-U/V	"Above 50"	"Above 50"
Westerville City	ITBS	"Above 50"	"Above 50"
Hamilton City	ITBS-G/H	"Above 50"	"Above 50"

Source: Telephone Survey, Spring 1989

TEST SECURITY IN OHIO

Ohio does not administer any statewide tests. They will require districts to use a state approved norm-referenced test beginning in the 1989-1990 school year, and the state will collect the data. There will be stringent state requirements concerning test construction and norming for the tests the state will allow the districts to use.

However, at this time, no state regulations prohibit teachers from looking at norm-referenced test booklets, prohibit teachers from obtaining the test booklets before the day of testing, require the booklets to be sealed, or mandate fall testing. In addition, at this time, testing will not be routinely monitored by state officials, special education students will not be specifically included, outside test proctors will not be required, test questions are not required to be rotated every year, and answer sheets will not necessarily be scanned for suspicious erasures or routinely analyzed for cluster variance.

OKLAHOMA *MARCH 1989*

THE METROPOLITAN
ACHIEVEMENT TEST
FORM M 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	43,209	61	62	65	64	66%	468/569 (82%)
7	38,178	58	59	60	61	61%	405/552 (73%)
10	35,511	53	57	53	56	54%	256/448 (57%)

Reporting method: median national individual percentiles
Source: The Oklahoma Report: An Update on Education 1989

TEST SECURITY IN OKLAHOMA

Oklahoma will administer a new test in 1989-1990, and they have recently adopted new security measures for their entire testing program. They plan to clearly forbid teachers to look at tests, they plan to deliver test booklets to the schools no earlier than 24 hours before testing, they plan to seal individual test booklets, and to test special education students unless their IEP specifically forbids testing. They will use outside test proctors to administer their norm-referenced test.

However, fall testing will not be required, the testing will not be routinely monitored by state officials, test questions will not be rotated every year, and answer sheets will not be routinely scanned for suspicious erasures or analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

OREGON* *SPRING 1988*

VARIOUS NATIONALLY NORMED TESTS
STANFORD, CALIFORNIA, CTBS AND IOWA MOST COMMON

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	30,323	61	N/A	62	N/A	78%	156/222(70%)
5	30,088	65	N/A	62	N/A	86%	159/217(73%)

Reporting method: mean national individual percentiles
Source: Summary of Local Achievement Test Results Grades 3 and 5 May 1989

TEST SECURITY IN OREGON

The statewide test, the Oregon Assessment Program, tests a sample of children from different schools every two years, but it is no longer nationally normed. The Oregon Assessment Program test uses new questions every time they test, teachers are handed the test just prior to administration, and special education students are included unless their IEP specifically forbids testing.

Oregon collected the above norm-referenced results from its districts, but there are no state regulations in Oregon concerning the ethical use of these tests in the districts. Individual school districts may establish their own test security, test procedure and test exclusion policies, but none are required by state law.

At the present time, there are no state regulations that forbid teachers from looking at the norm-referenced tests administered in local school districts, that forbid teachers from obtaining the test booklets before the day of testing, that require the booklets to be sealed, that requires fall testing, or that requires testing to be routinely monitored by state officials. In addition, special education students are not specifically required to be included in testing, outside test proctors are not required, test questions do not have to be rotated every year, and answer sheets are not necessarily scanned for suspicious erasures or routinely analyzed for cluster variance.

PENNSYLVANIA* *MARCH 1988* TESTING FOR ESSENTIAL LEARNING AND LITERACY SKILLS (TELLS) EQUATED WITH ITBS FORM ? 1978 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	117,948	78.9	N/A	86.2	N/A	N/A	N/A
5	112,366	79.8	N/A	82.3	N/A	N/A	N/A
8	115,290	80.1	N/A	79.2	N/A	N/A	N/A

Reporting method: mean percent of items correct (all scores except eighth grade reading are above the national mean percent correct).

Source: The Pennsylvania Testing and Assessment Program: 1987-88 Statewide Test Results

TEST SECURITY IN PENNSYLVANIA

Pennsylvania rotates most questions every year on the TELLS test. They will change all TELLS questions for the 1989-1990 school year. They also shrink-wrap their test booklets which are delivered to schools the day before testing, handed to the teacher on the day of testing, and answer sheets are scanned for suspicious erasures. Pennsylvania ia planning to analyze answer sheets for cluster variance.

However, teachers are allowed to look at the TELLS test booklets, fall testing is not required, testing is not routinely monitored by state officials, special education students are not generally tested unless their IEP indicates they are mainstreamed more than 50 percent of the time in either reading or math, and outside test proctors are not routinely used to administer the test.

There are no state regulations governing test security, test administration, or testing ethics for the commercial norm-referenced tests used independently in local Pennsylvania school districts.

RHODE ISLAND*

MARCH 1988 THE METROPOLITAN ACHIEVEMENT TEST FORM L 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	8,911	65	67	66	68	69%	36/38 (95%)
6	8,270	67	67	67	68	69%	35/37 (95%)
8	7,998	61	59	58	60	58%	30/34 (88%)
10	7,915	55	52	52	55	55%	25/31 (81%)

Reporting method: mean individual national percentiles
Source: Rhode Island State Assessment Program 1987-88.

TEST SECURITY IN RHODE ISLAND

Rhode Island administers the Metropolitan Achievement Test statewide. State officials monitor some test sessions, and they also require that special education students be included unless they spend most of day in special classes.

However, teachers are allowed to read test booklets because "the publisher's examiners manual recommends it." Teachers may obtain the Metropolitan test booklets before the day of testing, the booklets are not sealed, fall testing is not required, outside test proctors are not routinely used, test questions have not been rotated every year, and answer sheets are not routinely scanned for suspicious erasures or analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

SOUTH CAROLINA* *MARCH 1989* THE COMPREHENSIVE TEST OF BASIC SKILLS FORM U 1981 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
4	46,706	57.1	64.4	69.6	62.0	64.3%	81/92(88%)
5	45,047	51.6	59.3	66.9	55.2	55.5%	51/92(55%)
7	44,589	52.8	63.0	66.6	58.4	60.4%	71/92(77%)
9	47,676	49.1	58.3	60.2	54.2	53.8%	50/93(54%)
11	36,566	45.4	64.2	61.3	56.1	54.8%	48/93(52%)

Reporting method: median individual national percentiles

Source: South Carolina Statewide Testing Program, 1989 Summary Report

TEST SECURITY IN SOUTH CAROLINA

South Carolina also administers a graduation exam and a criterion referenced test, both of which have significant security measures. Teachers are not allowed to look at either of these two test booklets, teachers may not obtain booklets before the day of testing, the graduation test booklets are sealed, testing is routinely monitored by state officials, special education students are generally included in all tests used in South Carolina unless their IEP recommends against testing, outside test proctors administer the graduation exam, and most test questions are rotated every year on the criterion referenced test.

South Carolina has administered the CTBS-U form for six years in a row. They plan to administer a new form of the Stanford Achievement Test in the 1989-1990 school year. State law provides for a fine of $1,000 and imprisonment for up to 90 days for breeches of test security. Additional test security regulations are being considered for South Carolina's norm-referenced testing program.

Unlike their other two tests, teachers are allowed to look at CTBS test booklets, teachers may obtain CTBS test booklets before the day of testing, the booklets are not sealed, fall testing is not required, and CTBS testing is not routinely monitored by state officials. Outside test proctors are not routinely used to administer the CTBS, test questions have not been rotated every year, and CTBS answer sheets have not been routinely scanned for suspicious erasures or routinely analyzed for cluster variance.

There are no state regulations that govern test security and test administration for norm-referenced testing done independently in the local school districts.

SOUTH DAKOTA

APRIL 1989

THE STANFORD ACHIEVEMENT TEST
FORM J 1988 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
4	10,955	59	59	58	60	61%	N/A
8	9,231	57	58	62	58	60%	N/A
11	8,274	56	55	60	59	57%	N/A

Reporting method: mean individual national percentiles
Source: Spring 1989 South Dakota State Testing Program Results.

TEST SECURITY IN SOUTH DAKOTA

South Dakota administered a new form of the Stanford Achievement Test this year, and scores dropped significantly compared to the 1982 Stanford. South Dakota does test special education students unless their IEP specifically excludes them.

However, teachers are allowed to look at the test booklets, teachers may obtain the test booklets before the day of testing, the booklets are not sealed, fall testing is not required, the testing is not routinely monitored by state officials, outside test proctors are not routinely used, and test questions have not been rotated every year. In addition, answer sheets are not routinely scanned for suspicious erasures, or analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

South Dakota is one of the eleven states that is not planning to participate in the 1990 National Assessment of Educational Progress state-by-state evaluation.

TENNESSEE* *SPRING 1989*

THE STANFORD
ACHIEVEMENT TEST
FORM ? 1982 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50*	% DISTRICTS > OR = 50 **
2	63,131	61	69	77	N/A	61%	138/139(99%)
5	60,498	55	58	65	N/A	59%	97/139(70%)
7	59,123	50	57	59	N/A	56%	86/134(64%)
9	60,374	44	52	52	N/A	51%	76/125(61%)
12	49,148	55	58	55	N/A	56%	106/121(88%)

Reporting method: mean individual national percentiles

Source: Tennessee Student Test Results - 1988-89

* estimated from stanines in language

** in language

TEST SECURITY IN TENNESSEE

Tennessee will administer the new CTBS-4 this year and plans to rotate test forms every year. Testing generally includes special education students unless their IEP prohibits testing, and answer sheets are analyzed for cluster variance. Additional test security policies are now being considered in Tennessee.

Currently, teachers are allowed to look at the test booklets, teachers may obtain the test booklets before the day of testing, the booklets are not sealed, fall testing is not required, testing is not routinely monitored by state officials, outside test proctors are not routinely used, and answer sheets are not routinely scanned for suspicious erasures.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

Tennessee is one of the eleven states that is not planning to participate in the 1990 National Assessment of Educational Progress state-by-state evaluation.

TEXAS* *FEBRUARY 1988* TEXAS EDUCATIONAL ASSESSMENT OF MINIMUM SKILLS EQUATED WITH METROPOLITAN ACHIEVEMENT TEST FORM L 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	WRITING	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
1	254,099	819	845	860	N/A	N/A	N/A
3	235,812	809	765	847	N/A	N/A	N/A
5	223,758	808(51)	794(61)	819(57)	N/A	N/A	N/A
7	223,366	790(43)	770(59)	831(51)	N/A	N/A	N/A
9	227,609	788(57)	742(56)	798(60)	N/A	N/A	N/A

Reporting method: mean scaled score (national percentile rank of representative statewide sample in parenthesis)

Source: Student Performance Results 1987-1988

TEST SECURITY IN TEXAS

The Texas Educational Assessment of Minimum Skills (TEAMS) test is sealed, testing is monitored by state officials, special education students are tested unless their IEP forbids it, and approximately 50% of the TEAMS questions have been new every year. New ethics legislation in Texas addresses both the TEAMS and the norm-referenced tests used independently in the districts.

Under this law, all "group achievement tests" administered anywhere in Texas will be required to be secure, teachers will not have access to test materials until the day of testing, norms will be required to be current and accurate, tests will be required to be delivered to schools no earlier than a week before testing, educators will be required to sign a security oath, and tests will be required to be administered in the early fall.

However, regulations will not require answer sheets to be routinely scanned for suspicious erasures or analyzed for cluster variance.

Publishers will be held liable for punitive damages for selling deceptive tests in Texas. Publishers and test scoring companies will be required to submit district scores to the state.

UTAH *APRIL 1987* THE COMPREHENSIVE TEST OF BASIC SKILLS FORM U 1981 NATIONAL NORMS

GRADE	NUMBER TESTED*	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
5	4,930	61	55	65	N/A	N/A	N/A
11	2,582	63	65	62	N/A	N/A	N/A

Reporting method: median national individual percentiles

Source: Utah Statewide Educational Assessment General Report 1987

*Sample of students tested

TEST SECURITY IN UTAH

Utah administers their statewide assessment every three years to a randomly stratified group of schools. They use outside test proctors to administer the tests and clearly forbid teachers to look at the test. They test special education students unless their IEP forbids it, randomly monitor their assessment, and analyze answer sheets for cluster variance.

Unfortunately, Utah has no similar state security measures or code of testing ethics for the norm-referenced tests that are administered independently in the local districts. Individual school districts in Utah may establish their own test security, test procedure and test exclusion policies, but none are required by state law.

Utah is one of the eleven states that is not planning to participate in the 1990 National Assessment of Educational Progress state by state evaluation.

VERMONT

SPRING 1989 VARIOUS NATIONALLY NORMED TESTS

DISTRICT	TEST	ELEMENTARY	SECONDARY
Burlington	SAT-7	"Above 50"	"Above 50"
South Burlington	MAT-6	"Above 50"	"Above 50"
Southwest Vermont	CAT	"Above 50"	"Above 50"
Rutland City	SAT-7	"Above 50"	"Above 50"
Windom S.E.	CAT-E	"Above 50"	"Above 50"
Berry City	CTBS-U	"Above 50"	"Above 50"
Springfield	CAT	"Above 50"	"Above 50"
Montpelier	SAT	"Above 50"	"Above 50"
Saint Johnsbury	SAT-E	"Above 50"	"Above 50"

Source: Telephone Survey, Spring 1989

TEST SECURITY IN VERMONT

Vermont does not administer any statewide tests. They are planning to administer a statewide test in 1990-91 with "links to national data."

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local Vermont school districts. No state regulations prevent teachers from looking at norm-referenced test booklets, teachers may obtain the test booklets before the day of testing, booklets are not required to be sealed, fall testing is not required, testing is not routinely monitored by state officials, special education students are not specifically included, outside test proctors are not required, test questions do not have to be rotated every year, and answer sheets are not necessarily scanned for suspicious erasures or routinely analyzed for cluster variance.

Vermont is one of the eleven states that is not planning to participate in the 1990 National Assessment of Educational Progress state by state evaluation.

VIRGINIA* APRIL 1989 THE IOWA TEST OF BASIC SKILLS FORM G 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50*	% DISTRICTS > OR = 50*
4	71,500	54	60	60	N/A	61%	99/133(72%)
8	66,500	54	57	56	N/A	58%	89/132(67%)
					TEST OF ACHIEVEMENT AND PROFICIENCY		
11	59,500	57	61	56	N/A	58%	116/131(88%)

Reporting method: mean individual national percentiles

Source: Virginia State Assessment Program Results 1988-89

* In language

TEST SECURITY IN VIRGINIA

This was the first year the ITBS was administered statewide in Virginia. Virginia plans to administer a criterion-referenced test to sixth graders beginning in the 1989-90 school year. Students will be required to pass this test during the sixth, seventh, or eighth grade in order to be promoted to the ninth grade. According to state officials, this criterion-referenced test will be "secure," and new test forms will be used each year.

Virginia is revising its test security measures for norm-referenced tests. Current policy requires that teachers be given the Iowa test booklets no more than 24 hours before testing, and that special education students be tested unless their IEP specifically excludes them. In addition, state officials routinely analyze Iowa answer sheets for cluster variance.

However, at the present time, teachers are allowed to look at the test booklets, booklets are not sealed, and fall testing is not required. In addition, testing is not routinely monitored by state officials, outside test proctors are not routinely used, test questions have not been rotated every year, nor are answer sheets routinely scanned for suspicious erasures.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

WASHINGTON* *OCTOBER 1988*

THE METROPOLITAN ACHIEVEMENT TEST
FORM L 1985 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
4	58,148	56	53	54	57	57%	165/262(63%)
8	50,407	60	55	56	59	59%	171/249(69%)
10	47,340	56	48	57	54	55%	139/237(59%)

Reporting method: percentile equivalent of median national normal curve equivalent
Source: Washington Statewide Assessment Grades 4 8 10: Fall 1988

TEST SECURITY IN WASHINGTON

Washington administers the MAT-6 in early October and includes special education students unless their IEP specifically prohibits them from being tested. Packages of test booklets are shrunk wrapped for delivery to schools.

State regulations hold that it is a misdemeanor for educators to reveal examination answers to students. However, no other regulations govern test security and test administration for the norm-referenced testing done independently in the local school districts.

In Washington, teachers are allowed to look at the test booklets "because the publisher's examiners manual recommends it", teachers may obtain the test booklets before the day of testing, booklets are not sealed, testing is not routinely monitored by state officials, outside test proctors are not routinely used, test questions have not been rotated every year, and answer sheets are not routinely scanned for suspicious erasures or routinely analyzed for cluster variance.

Washington is one of the eleven states that is not planning to participate in the 1990 National Assessment of Educational Progress state-by-state evaluation.

WEST VIRGINIA

MARCH 1988 THE COMPREHENSIVE TEST OF BASIC SKILLS FORM U 1981 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50	% DISTRICTS > OR = 50
3	22,291	64	73	57	68	72.8%	55/55(100%)
6	22,504	58	63	67	62	64.7%	52/55(95%)
9	22,836	49	58	56	52	53.8%	36/55(65%)
11	22,294	52	59	59	58	56.1%	43/55(78%)

Reporting method: mean individual national percentiles

Source: West Virginia State-County Testing Program: 1987-88 Results.

TEST SECURITY IN WEST VIRGINIA

West Virginia has administered the 1981 CTBS-U for five years and plans to administer the CTBS-U again this year. They plan to administer a state-developed criterion referenced test in the 1990-1991 school year.

Teachers are allowed to look at the CTBS test booklets, teachers may obtain the booklets before the day of testing, booklets are not sealed, fall testing is not required, testing is not routinely monitored by state officials, special education students are generally excluded unless their IEP recommends testing, outside test proctors are not routinely used, test questions have not been rotated every year, and answer sheets are not routinely scanned for suspicious erasures or analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

WISCONSIN* *SPRING 1987/88* THE COMPREHENSIVE TEST OF BASIC SKILLS FORM U 1981 NATIONAL NORMS

GRADE	NUMBER TESTED	READING	LANGUAGE	MATH	TOTAL BATTERY	% STUDENTS > OR = 50*	% DISTRICTS > OR = 50
4	3000	74	70	70	N/A	73%	N/A
8	3000	67	66	74	N/A	69%	N/A
11	3000	61	N/A	74	N/A	69%	N/A

Reporting method: median individual national percentiles
Source: Wisconsin Department of Public Instruction
* in language

TEST SECURITY IN WISCONSIN

Wisconsin stopped administering the CTBS in 1988. The Wisconsin Department of Public Instruction now administers a state-developed reading test at third grade. The state tells teachers not to read the test except as needed for administration, teachers are provided with the test booklets shortly before testing, the packages of test booklets are sealed, the state randomly monitors testing, and develops a new test annually.

However, the testing of special education students is at the discretion of local school districts (the state provides voluntary test exclusion "guidelines"), outside test proctors are not routinely used, fall testing is not required, and answer sheets are not routinely scanned for suspicious erasures or analyzed for cluster variance.

There are no state regulations that govern test security and test administration for the norm-referenced testing done independently in the local school districts.

WYOMING *SPRING 1989*

VARIOUS NATIONALLY NORMED TESTS

DISTRICT	TEST	ELEMENTARY	SECONDARY
Cheyenne	SRA-M	"Above 50"	"Above 50"
Natrona	SAT-7	"Above 50"	"Above 50"
Campbell	MAT-6	"Above 50"	"Above 50"
Sweetwater #1	SRA	"Above 50"	"Above 50"
Albany	MAT-6	"Above 50"	"Above 50"
Sheridan #2	SRA	"Above 50"	"Above 50"
Uinta	SRA	"Above 50"	"Above 50"

Source: Telephone Survey, Spring 1989

TEST SECURITY IN WYOMING

Wyoming administers no statewide test of any kind. The Wyoming State Department of Education will start collecting local district's test results next year.

However, there are no state regulations that govern test security and test administration for the norm-referenced testing that will be compiled by the state. No state regulations will prohibit teachers from looking at norm-referenced test booklets, teachers may obtain the test booklets before the day of testing, booklets are not required to be sealed, fall testing is not required, testing is not routinely monitored by state officials, special education students are not specifically included, outside test proctors are not required, test questions do not have to be rotated every year, and answer sheets will not necessarily be scanned for suspicious erasures or routinely analyzed for cluster variance.

APPENDIX II

SURVEY OF TEST SECURITY PRACTICES IN THE 50 STATES

This appendix contains a table that lists state-by-state test security practices in the fifty states. The practices listed in Table I are for any nationally normed tests administered or compiled by that state's officials. This includes off-the-shelf norm-referenced tests, state developed tests normed by concurrent administration of a commercial test, tests normed within the last three years by an equating study, tests normed with NAEP data, and tests administered locally but compiled to any extent by state officials. If the state does not administer or compile the scores for nationally-normed tests, then the table refers to statewide criterion referenced testing. If the state has no statewide testing at all, then the table refers to state security policies for the norm-referenced tests used in that state's local school districts.

The ten states with asterisks have not recently administered any nationally normed tests, nor have they collected this data from their districts. Six of these ten states administer criterion referenced tests, and the data in Table I refers to their security policies for these tests. Four states, Montana, Nebraska, Vermont, and Wyoming, have no statewide testing, nor do they compile local district's scores. For these four states, the table refers to state regulations for norm-referenced tests administered independently in their local school districts.

In one sense, the table is misleading because it shows states like New York, California, Florida, and Pennsylvania as all having significant test security measures. However, these are for their state administered tests. Like the four states mentioned above, they have no test security regulations for the norm-referenced tests used in their local school districts. Only Alabama, Texas, Ohio, and Washington have testing regulations that cover the "Lake Wobegon" tests administered independently in their local school districts.

EXPLANATION OF ABBREVIATIONS

STA: STAtes are listed in alphabetical order.

SEC: SECurity policy. Does the state have a written test security policy?

PRO: PROcedure policy. Does the state have a written procedure policy?

TNL: Told Not to Look. Does the state have a written policy clearly telling teachers not to look at the test, except as needed at the time of administration.

TDT: Tests on the Day of Testing. Does the state have a written policy stating that the teachers are to receive the tests no earlier than the day of testing.

SEA: SEAled. Are the test booklets individually sealed or are they shrunk

wrapped in class sized packets when delivered to the districts every year?

FAL: FALL. Are tests administered in early fall?

RMT: Random Monitoring of Testing. Does the state do unannounced monitoring of test administration and test security?

TSE: Test Special Education. Does the state's test procedure policies state that special education students are tested unless their IEP specifically prohibits testing?

PRC: PRoCtors. Are outside proctors used to supervise the testing?

ROT: ROTation of equivalent test questions. Does the state rotate test questions either by rotating forms, by changing test questions every year, by administering multiple matrix forms of the same test, or by administering the test to a different stratified sample of schools every year?

SES: Suspicious Erasure Scanning. Does the state scan the answer sheets for suspicious numbers of wrong answers erased and changed to right ones?

CVA: Cluster Variance Analysis. Does the state routinely analyze answer sheets for cluster variance?

INV: INVestigations. How many investigations of irregular testing practices has the state processed, either formally or informally?

CON: CONfirmed. How many of these investigations have been confirmed?

EXPLANATION OF SYMBOLS

X The state has this particular security policy in place

O The state does not have this particular security policy in place

P The state is planning to institute this security policy

C The state claims to have this particular security policy but we have not been able to confirm it in their written security policies

TEST SECURITY PRACTICES IN THE 50 STATES

STA	SEC	PRO	TNL	TDT	SEA	FAL	RMT	TSE	PRC	ROT	SES	CVA	INV	CON
AL	X	X	P	P	O	O	X	X	O	O	O	O	6-10	4
AK	X	X	O	O	O	P	O	P	O	O	O	O	1-2	0
AZ	X	X	O	X	O	O	X	O	O	O	O	O	20+	12
AR	X	X	P	X	P	P	X	O	O	P	P	O	3-5	3
CA	X	X	X	X	O	O	X	X	X	X	X	O	>70	70
CO	X	X	O	X	O	X	X	X	O	X	O	O	1-2	0
CT	X	X	O	O	O	X	X	X	O	X	O	O	11-20	5
DE	X	X	O	O	O	O	X	X	O	X	O	O	1-2	0
FL	X	X	X	X	X	O	O	X	O	X	X	O	3-5	3-5
GA	X	X	X	X	O	O	X	X	X	O	O	O	2	1
HI	P	X	O	O	O	O	P	O	O	O	O	O	3-5	0
ID	X	X	O	X	O	O	O	X	O	O	O	O	3-5	2

STA	SEC	PRO	TNL	TDT	SEA	FAL	RMT	TSE	PRC	ROT	SES	CVA	INV	CON
IL	X	X	O	O	X	O	O	O	O	X	O	O	0	0
IN	X	X	O	O	O	O	X	X	O	O	O	O	0	0
LA	X	X	O	O	O	X	O	O	O	X	O	O	3-4	0
KS*	X	X	O	O	X	O	O	X	O	X	O	O	2-3	2
KY	X	X	O	O	O	O	O	X	O	O	O	O	3-5	2
LA	X	X	O	O	O	O	O	X	O	O	O	O	3	3
ME	X	X	O	X	X	X	X	X	X	X	X	X	5	1
MD	X	X	X	X	P	X	P	X	O	P	P	P	6-10	3
MA	X	X	O	O	X	O	O	X	O	X	X	O	5	3
MI*	X	X	O	O	X	X	X	X	O	P	X	X	5-10	2
MN*	X	X	O	O	O	O	O	X	X	X	O	O	0	0
MS	X	X	P	X	O	O	X	X	O	O	P	P	20+	?
MO	X	X	P	P	X	O	O	X	O	X	O	O	3-5	0
MT*	O	O	O	O	O	O	O	O	O	O	O	O	0	0
NE*	O	O	O	O	O	O	O	O	O	O	O	O	0	0
NV	X	X	O	O	O	O	O	X	O	O	O	O	1	1
NH	O	X	C	C	C	X	O	O	O	O	O	O	0	0
NJ*	X	X	O	O	O	O	O	X	O	O	O	O	11-20	4
NM	O	X	O	O	O	O	O	O	O	O	O	O	0	0
NY*	X	X	X	X	X	O	X	X	X	X	O	O	4-6	4
NC	C	C	C	C	C	O	C	C	C	O	O	O	6-10	4+
ND	O	O	O	O	O	O	O	O	O	O	O	O	0	0
OH*	X	X	O	O	O	O	O	O	O	O	O	O	0	0
OK	X	X	P	P	P	O	O	P	X	O	O	O	2	2
OR	X	X	O	X	O	O	O	X	O	X	O	O	0	0
PA	X	X	O	X	X	O	O	O	O	X	X	P	8	3
RI	X	X	O	O	O	O	X	X	O	O	O	O	0	0
SC	X	X	O	O	O	O	O	X	O	O	P	O	0	0
SD	X	X	O	O	O	O	O	X	O	O	O	O	0	0
TN	X	X	O	O	O	O	O	X	O	P	O	X	10-15	6
TX	X	X	P	P	X	P	X	X	O	X	O	O	>20	13
UT	X	X	X	X	O	O	X	X	X	X	O	X	0	0
VT*	O	O	O	O	O	O	O	O	O	O	O	O	0	0
VA	X	X	O	X	O	O	O	X	O	O	O	X	>20	4
WA	X	X	O	O	X	X	O	X	O	O	O	O	3-5	3
WV	X	X	O	O	O	O	O	O	O	O	O	O	3-5	3
WI	X	X	X	X	X	O	X	O	O	X	O	O	0	0
WY*	O	O	O	O	O	O	O	O	O	O	O	O	0	0

APPENDIX III

ARE AMERICAN SCHOOLS IMPROVING?

Publishers and school officials are quick to point out that scores on "Lake Wobegon" achievement tests have improved because student achievement has improved. However, why, as recently stated by the U.S. Department of Education, "have other national and international assessment programs not reported the kind of high achievement" found on norm-referenced achievement tests (12)? In Appendix III, we list some of those other indicators. **If the high scores on the "Lake Wobegon" tests are the result of steady improvement in American schools, why do national achievement tests show little or no improvement?**

College Entrance Scores Decline

College Board scores in the United States have declined three percent over the last 15 years, from 937 in 1972 to 904 in 1988 (1,2). ACT scores have fallen during the same time period, even though the ACT tests different students from different states with a very different test. The ACT national average was 19.1 in 1972 and had fallen to 18.8 by 1988 (1,2). America's graduation rate has fallen, too, from 76.9 percent in 1972 to 71.1 percent in 1987 (34,42). All three indicators fell at the same time that "Lake Wobegon" achievement scores reached all time highs. Even the recent modest gains on college entrance scores (two percent over six years), do not correspond to the dramatic gains on "Lake Wobegon" tests in the same time period (10).

Military Tests Show No Improvement

The United States Military tests more than one million high school students every year on basic reading and math skills. The Armed Forces Vocational Aptitude Battery (ASVAB) cannot document any improvements in either basic math or basic verbal skills. Both the math and verbal subtests of ASVAB were normed in 1980 with a standard score of 50; by 1986 the verbal composite was 46.4, and the math composite was 48 (43). Although the 1980 ASVAB norms were based on the general population, ASVAB has not confirmed the dramatic gains in American achievement that "Lake Wobegon" tests claim to have documented.

Educators often contend that ASVAB and college entrance tests are not useful for tracking American achievement because they are self-selective; that is, they only test a certain self-selected sub-group of students, not all the students in the system. ASVAB only tests those students considering joining the Armed Services, and college entrance tests only tests those students considering college. However, national surveys of young adult illiteracy and biennial national sampling of student achievement corroborate the decline in ASVAB and college entrance scores, and confirm the U.S. Department of Education's contention that other indicators of

American student achievement do not show the rapid progress shown on "Lake Wobegon" tests.

National Reading Skills Flat

The federal government conducts a national survey of reading skills every two years called the National Assessment of Educational Progress (NAEP). American elementary reading scores (age 9) were flat from 1980 to 1986 (5,44). Scores for other grade levels and other subjects have either remained flat or shown slight increases in the 1980's; none have shown the dramatic increases reported on the "Lake Wobegon" tests. NAEP scores for 1988 will also be flat (45).

Who's Illiterate

The 1985 National Assessment of Educational Progress (NAEP) conducted a survey of young adult literacy and found that 40 percent of young Americans educated in the last 15 years could not read a map, use an almanac, or balance a checkbook (4). The 1985 NAEP illiteracy figure was almost 80 percent for young Black Americans. **That means the vast majority of Black Americans educated in American public schools in the last 15 years cannot read a map, balance a checkbook, use an almanac, or understand a bus schedule; they are functionally illiterate.**

Are Lake Wobegon" Gains Real?

However, the publishers contend that American elementary achievement has improved dramatically in recent years according to their norm-referenced tests. It is true that elementary state assessments have recorded steady, often dramatic, achievement gains in the 1980s in every state that uses the same test questions year after year. In the second edition of *How All 50 States Are Above The National Average*, we looked again at the volumes of test scores we collected from all over the country. All states show steady achievement gains, regardless of school quality, if they used "Lake Wobegon" tests without rotating test questions in some manner (46).

Not only do other indicators of achievement fail to support the dramatic gains seen on "Lake Wobegon" tests, but school officials admit they cannot detect any dramatic gains in secondary achievement over the last 15 years. Why not? If the elementary "Lake Wobegon" gains are real, one assumes that such gains will eventually show up in secondary schools. Do American teenagers just forget how to read once they reach secondary school?

Explanation of Appendix III

This appendix lists state-by-state rankings for three independent education statistics. Unfortunately, no perfect indicators of state-by-state achievement exist. All the ones that currently exist suffer from severe limitations, and none were designed to evaluate public education.

All the standardized tests listed in this appendix use strict security. None of them allow teachers or administrators to see test questions in advance of testing.

They either rotate equivalent test questions every year or use outside examiners. Needless to say, none of them allow more than one-half of the students tested every year to be above average.

However, the three tests listed in this appendix do not test all students in the district; they are self-selective, as explained above. In addition, all of the tests we cite are high school assessments because no elementary achievement results are currently available on a state-by-state basis, other than the "Lake Wobegon" tests.

Table I

College entrance scores are one method of comparing local achievement to national achievement. They are the best available method of answering this question: how do the college bound students in my district compare to other college bound students in the country? Keep in mind that the tests are self-selective, that two different tests are used in the country, and that the participation rate can effect your district's scores.

Table I lists the most recent scores of the 22 "SAT states" and the "28 ACT states." States are classified as ACT or SAT, depending on which test is required for admission to that state's university. All 50 states have been combined into one table because citizens often misunderstand separate rankings.

For example, Mississippi is listed 28th for ACT scores on the U.S. Department of Education's "Wall Chart." This means Mississippi ranks last on college entrance scores, not 28th, because there are only 28 "ACT states," that is, 28 states require the ACT for admission to their state colleges and universities.

Participation rates on college entrance tests make state-by-state comparisons difficult because larger participation rates generally will lower the state's score. In addition, the test is self-selecting in that only those students considering college take the test. With these limitations in mind, one can compare a state's college entrance score with the national average for that test.

Table II

Every year the Armed Services Vocational Aptitude Battery is administered to the more than one million high school students who are considering joining the Armed Services. The test is a norm-referenced test that is administered under strict security by military personnel. Table II lists the state-by-state scores, as well as the number of students taking the test in each state.

Table III

Table III lists the percent of childhood poverty in each state, as published by the U.S. Department of Education. Socioeconomic factors, not school quality, is the strongest determinant of school achievement.

It is striking how three very different measurements leave similar impressions of state rankings. Look, for example, where West Virginia ranks on all three

tables. Then compare the relative positions that Arkansas places on the three tables. With the exception of six states, the relative rankings of the 50 states are roughly similar on all three tables.

Compare West Virginia's "Lake Wobegon" test scores in Appendix I with the three tables in this appendix. On the most recent administration of the Comprehensive Test of Basic Skills, West Virginia ranked above the national norm at all grade levels tested.

In contrast, Tables I through III show that West Virginia has the the fourth lowest college entrance scores (ACT) and the ninth lowest ASVAB scores in the nation. In addition, they rank well below the national average in childhood poverty.

Arkansas is well above average on the Metropolitan Achievement Test. However, Appendix II shows that Arkansas ranks well below average on all the tables in this appendix.

It is apparent that many states' "Lake Wobegon" achievement scores are strikingly different from other indicators of achievement. We are forced to use imperfect indicators; none of which were designed to make comparisons. However, it is apparent that large discrepancies exist for many states between their "Lake Wobegon" test results and the other indicators of their achievement listed in this appendix.

School board members are urged to obtain district reports of college entrance scores by writing to one of the following addresses. If your school district is in one of the 28 ACT states write to:

American College Testing
P.O Box 168
Iowa City, Iowa 52243

If your school district is in one of the 22 SAT states write to:

The College Board
45 Columbus Avenue
New York, New York 10023

School board members are also urged to obtain their districts' Armed Services Vocational Aptitude Battery Scores from the ASVAB representative in their area. Your local military recruiting station will supply you with their name.

TABLE I

1988 COLLEGE ENTRANCE SCORES

RANK	STATE	SCORE		RANK	STATE	SCORE	
1]	Iowa	20.3	(ACT)	26]	Nevada	19.0	(ACT)
2]	New Hampshire	933	(SAT)	27]	New Jersey	893	(SAT)
3]	Wisconsin	20.2	(ACT)	28]	Florida	890	(SAT)
4]	Oregon	923	(SAT)	29]	Utah	18.9	(ACT)
5]	Minnesota	19.9	(ACT)	30]	Illinois	18.9	(ACT)
6]	Montana	19.9	(ACT)	31]	New York	889	(SAT)
7]	Vermont	909	(SAT)	32]	Hawaii	888	(SAT)
8]	Maryland	908	(SAT)	33]	Pennsylvania	886	(SAT)
9]	Connecticut	908	(SAT)	34]	Michigan	18.8	(ACT)
10]	California	908	(SAT)	35]	Texas	879	(SAT)
11]	South Dakota	19.8	(ACT)	36]	North Dakota	18.7	(ACT)
12]	Nebraska	19.8	(ACT)	37]	Indiana	870	(SAT)
13]	Colorado	19.7	(ACT)	38]	Alaska	18.4	(ACT)
14]	Massachusetts	906	(SAT)	39]	Kentucky	18.2	(ACT)
15]	Wyoming	19.5	(ACT)	40]	Alabama	18.1	(ACT)
16]	Arizona	19.3	(ACT)	41]	Georgia	848	(SAT)
17]	Ohio	19.3	(ACT)	42]	Tennessee	18.0	(ACT)
18]	Idaho	19.3	(ACT)	43]	New Mexico	18.0	(ACT)
19]	Virginia	902	(SAT)	44]	Oklahoma	18.0	(ACT)
20]	Rhode Island	900	(SAT)	45]	North Carolina	841	(SAT)
21]	Delaware	899	(SAT)	46]	Arkansas	17.9	(ACT)
22]	Kansas	19.1	(ACT)	47]	West Virginia	17.6	(ACT)
23]	Missouri	19.1	(ACT)	48]	South Carolina	838	(SAT)
24]	Maine	896	(SAT)	49]	Louisiana	17.1	(ACT)
25]	Washington	N/A		50]	Mississippi	16.2	(ACT)

1988 state ACT scores this year ranged from 16.2 to 20.3 with a national average of 18.8. SAT scores ranged from 838 to 933 with a national average of 904.

SOURCE: U.S. Department of Education, Office of Planning, Budget and Evaluation [1989]. *State Education Statistics, Student Performance, Resource Inputs, State Reforms and Population Characteristics 1982 and 1988.* Washington, D.C.

TABLE II

ARMED SERVICES VOCATIONAL APTITUDE BATTERY AVERAGES SCHOOL YEAR 1985-86

RANK	STATE	COMPOSITE	MATH	VERBAL	NO. TESTED
1]	Montana	51.72	53.44	50.44	6,049
2]	New Hampshire	51.62	53.14	50.48	3,688
3]	North Dakota	51.37	52.93	49.77	5,126
4]	Alaska	51.16	52.82	49.91	2,231
5]	Nebraska	51.14	53.38	49.59	9,035
6]	Minnesota	51.09	52.61	49.56	11,235
7]	Wisconsin	50.91	52.81	49.52	18,841
8]	Kansas	50.87	52.40	49.35	6,587
9]	Washington	50.86	51.99	49.94	18,594
10]	Iowa	50.77	52.62	49.25	10,564
11]	Idaho	50.69	51.54	49.76	7,009
12]	South Dakota	50.30	52.00	48.67	3,729
13]	Wyoming	49.81	51.18	48.92	2,345
14]	Massachusetts	49.69	51.21	48.70	13,023
15]	Utah	49.63	50.56	48.54	8,403
16]	Oregon	49.52	50.71	48.58	11,484
17]	Rhode Island	49.51	51.21	48.80	2,624
18]	Maine	49.47	50.87	48.48	5,093
19]	Vermont	49.39	50.66	48.20	1,816
20]	Colorado	49.13	51.02	48.35	9,806
21]	Nevada	49.03	50.57	48.04	2,684
22]	Pennsylvania	48.98	50.72	48.12	46,470
23]	Connecticut	48.91	50.62	47.96	5,997
24]	Ohio	48.86	50.59	47.85	41,000
25]	Michigan	48.79	50.75	47.48	27,563
26]	New York	48.67	51.51	47.31	33,473
27]	Florida	48.41	50.97	46.77	50,631
28]	Indiana	48.10	50.47	47.08	23,062
29]	Delaware	48.02	50.65	47.10	1,583
30]	New Jersey	47.91	50.21	46.84	15,548
31]	California	47.73	50.09	46.39	74,245
32]	Virginia	47.51	49.70	46.50	18,444
33]	Illinois	47.28	49.65	46.33	37,778
34]	Kentucky	47.25	49.17	46.14	20,743
35]	Maryland	47.11	49.51	46.00	10,738
36]	Missouri	46.82	49.04	45.75	37,294
37]	North Carolina	46.74	49.35	45.81	33,730
38]	Arkansas	46.63	48.83	45.82	21,241
39]	Arizona	46.41	48.54	45.66	15,837
40]	Alabama	46.40	48.97	45.02	35,576
41]	West Virginia	46.33	47.82	45.53	10,993
42]	Oklahoma	46.25	47.90	45.64	24,443
43]	Georgia	46.22	49.37	45.10	46,231
44]	Texas	46.01	48.78	44.93	78,100
45]	Tennessee	45.96	48.16	44.93	33,772

46] Louisiana	45.34	48.42	44.28	36,425
47] New Mexico	45.11	47.46	44.35	12,322
48] Hawaii	45.08	49.55	43.06	7,243
49] South Carolina	44.95	48.63	43.87	18,422
50] Mississippi	43.62	47.43	42.71	30,304
U.S. AVERAGE	**47.39**	**49.80**	**46.25**	**1,029,920**

SOURCE: *United States Department of Defense, Manpower Data Center, 1600 N Wilson Blvd., Arlington Virginia, 22209-2593.*

TABLE III

PERCENT POVERTY AGES 5-17

RANK STATE	PERCENT	RANK STATE	PERCENT
1] Wyoming	7.5	26] Idaho	13.4
2] New Hampshire	8.9	27] North Dakota	14.0
3] Nevada	9.4	28] Missouri	14.0
4] Minnesota	9.5	29] Illinois	14.1
5] Wisconsin	9.6	30] California	14.2
6] Utah	9.8	31] Virginia	14.4
7] Washington	10.3	32] Delaware	14.6
8] Indiana	10.3	33] Oklahoma	15.1
9] Connecticut	10.4	34] Maine	15.1
10] Kansas	10.7	35] Arizona	15.8
11] Oregon	10.8	36] Florida	17.7
12] Iowa	10.8	37] North Carolina	17.8
13] Colorado	10.8	38] New York	17.9
14] Alaska	11.4	39] West Virginia	18.2
15] Nebraska	11.6	40] Texas	18.4
16] Hawaii	11.7	41] South Dakota	19.4
17] Maryland	11.9	42] Tennessee	20.2
18] Ohio	12.2	43] Georgia	20.5
19] Massachusetts	12.3	44] South Carolina	20.7
20] Michigan	12.4	45] Kentucky	21.2
21] Rhode Island	12.6	46] New Mexico	21.7
22] Montana	12.7	47] Arkansas	22.7
23] Vermont	13.0	48] Louisiana	23.1
24] Pennsylvania	13.2	49] Alabama	23.1
25] New Jersey	13.3	50] Mississippi	30.4

U.S. AVERAGE 15.3

SOURCE: *Grant, V., [1987]. Center for Education Statistics. Fall Statistics of Public Schools and Digest of Educational Statistics. Department of Education, Washington D.C..*

REFERENCES

1) U.S. Department of Education, Office of Planning, Budget and Evaluation [1989]. *State Education Statistics, Student Performance, Resource Inputs, State Reforms and Population Characteristics 1982 and 1988.* Washington, D.C.

2) U.S. Department of Education, Office of Planning, Budget and Evaluation [1984]. *State Education Statistics, State Performance Outcomes, Resource Inputs and Population Characteristics 1972 and 1982.* Washington D.C.

3) Wall Street Journal [1989]. *The Wall Street Journal Reports, Education.* March 31, 1989.

4) Kirsch, I., and Jungeblut, A. [1986]. Literacy Profiles of Young Adults. Educational Testing Service, Princeton, NJ.

5) Applebee, A., Langer, J., and Mullis I. [1988]. *Who Reads Best? Factors Relating to Reading Achievement In Grades 3, 7, and 11.* Educational Testing Service, Princeton NJ.

6) Lapointe, A., Mead N., and Philips, G. [1989]. *A World of Differences.* Educational Testing Service, Princeton, NJ.

7) State of Alabama, Department of Education [1989]. *For Immediate Release.* 1989 Results of the Stanford Achievement Test. Montgomery, AL.

8) Office of Research, South Carolina Department of Education [1989]. *South Carolina Statewide Testing Program, 1989, Summary Results.* Columbia, SC.

9) Chicago Public Schools [1986]. *Comparison of Scores on the Iowa Test of Basic Skills for 1985-86, Grade Equivalent by Grade.* Chicago, IL.

10) Cannell, J. [1988]. *Nationally Normed Elementary Achievement Testing in America's Public Schools: How All 50 States Are Above the National Average.* Educational Measurement: Issues and Practice, 7(2).

11) Mitgang, L. [1987]. *Study Challenges Standardized Test Results.* The Associated Press, November 27, 1987.

12) Phillips, G. W., and Finn, C.E. [1988]. *The Lake Wobegon Effect: A Skeleton in the Testing Closet?* Educational Measurement: Issues and Practice, 7(2).

13) Rothman, R. [1989]. *"Lake Wobegon Effect" Real in Tests, Study Says.* Education Week, 8(28).

14) Linn, R., Graue, M., and Sanders, N. [1989]. *Comparing State and District Test Results to National Norms: Interpretation of Scoring "Above the National Average."* Paper presented at the annual meeting of the American Educational Research Association, March, 1989, San Francisco, CA.

15) Shepard, L. [1989]. *Inflated Test Score Gains: Is It Old Norms or Teaching to the Test?* Paper presented at the annual meeting of the American Educational Research Association, March, 1989, San Francisco, CA.

16) Koretz, D. [1988]. *Arriving in Lake Wobegon: Are standardized Achievement Tests Exaggerating Advancement and Distorting Instruction?* American Educator, 12(2).

17) Haney, W. [1985]. *Making Testing More Educational.* Educational Leadership, 43(2).

18) Perlman, C. [1985]. *Results of a Citywide Testing Program Audit in Chicago.* Paper presented at the annual meeting of the American Educational Research Association, March, 1985, Chicago, IL.

19) McGraw, C., and Wood T. [1988]. *Pressure for High Scores Blamed in CAP Test Cheating.* Los Angeles Times, September 18, 1988.

20) Ligon, G. [1985]. *Opportunity Knocked Out: Reducing Cheating by Teachers on Student Tests.* Paper presented at the annual meeting of the American Educational Research Association, April, 1985, Chicago, IL.

21) Mershon, C. [1986]. *How Parkrose Improved its Testing Program Using the Rasch Model.* Paper presented at the annual meeting of the National Council on Measurement in Education, April, 1986, San Francisco, CA.

22) Taylor, P. [1984]. *Texas Merit Pay Test Marked by Questions of Teacher Cheating.* Washington Post, June 9, 1984.

23) Wickendon, D. [1983]. *Merit Pay Won't Work.* The New Republic, November, 1983.

24) Chambers, M. [1980]. *Watch Is Under Way for Cheaters as Regents Exam Begins.* New York Times, June 20, 1980.

25) The New York Post [1989]. *Easy as PI. Stolen Regents Exams for Sale.* June 20, 1989.

26) Cohn, D. [1987]. *Five Fairfax Educators Disciplined over Coaching on Tests.* Washington Post, August 7, 1987.

27) Loepp, D. [1987]. *Some Teach to the Test.* The Daily Press, Newport News, VA, June 15, 1987.

28) Frary, R. and Olson, G. [1985]. *Detection of Coaching and Answer Copying on Standardized Tests.* Paper presented at the annual meeting of the National Council on Measurement in Education, April, 1985, Chicago, IL.

29) Savage, D. [1984]. *Scrutinize Students' Test Scores, and They Might Not Look So Rosy.* The American School Board Journal, August 1984.

30) Prell J., and Prell P. [1986]. *Improving Test Scores-Teaching Test Wiseness.* Research Bulletin, Phi Delta Kappa, November, 1986.

31) Somerby, R. [1988]. *Personal Communication.* February, 9, 1988, Baltimore, MD.

32) Ledonni, R. [1989]. *Personal Communication.* New York City Office of Research, Evaluation and Assessment. August 9, 1989.

33] Cannell, J. [1988]. *The Lake Wobegon Effect Revisited.* Educational Measurement: Issues and Practice 7(4).

34] U.S. Department of Education, Office of Educational Research and Improvement. [1987]. *Digest of Education Statistics, 1987.* Washington, D.C.

35) Mehrens, W. and Kaminski J. [1989]. *Methods for Improving Standardized Test Scores: Fruitful, Fruitless, or Fraudulent.* Educational Measurement: Issues and Practice 8(1).

36) The Riverside Publishing Company [1987]. *Test Resource Catalog.* Chicago, IL.

37) Science Research Associates [1988]. *Science Research Associates School Catalog.* Chicago, IL.

38) CTB/McGraw Hill [1985]. *CAT Parent Report [1-131-PR].* Monterey CA.

39) Delaware Department of Public Instruction [1989]. *Delaware Educational Assessment Program, Statewide Test Results, Summary Report.* Dover, DE.

40) West Virginia Department of Education [1987]. *Quick Facts About West Virginia's Public Schools.* Charleston, WV.

41) New Mexico State Department of Education [1986]. *New Mexico School District Profile 1985-1986 School Year.* Santa Fe, NM.

42) Grant, V. [1987]. Center for Education Statistics. *Fall Statistics of Public Schools and Digest of Educational Statistics.* Department of Education, Washington, D.C.

43) Sellman, W. [1987]. *Personal Communication.* The Office of the Assistant Secretary of Defense, Military Manpower and Personnel Policy, Washington, D.C.

44) Beaton E., et al. [1987]. *Final Draft, The NAEP 1985-86 Reading Anomaly; A Technical Report.* Educational Testing Service, Princeton, NJ.

45) Mullis, I. [1989]. *Personal Communication.* National Assessment of Educational Progress, Educational testing Service, Princeton, NJ.

46) Cannell, J. [1987b]. *Nationally Normed Elementary Achievement Testing in America's Public Schools: How All 50 States Are Above The National Average (2nd edition).* Friends for Education, Daniels, WV.

ABOUT THE AUTHOR

John Jacob Cannell M.D. was born in Washington, D.C. in 1948. He attended college at the University of Maryland and medical school at the University of North Carolina. He is president and founder of Friends for Education. He currently lives in Albuquerque, New Mexico with his wife, Ina, and their three children: Jacob, age 10, Travis, age 7 and Eliza, age 5 months.

Here are some of the responses to John Jacob Cannell's landmark report, "How All 50 States Are Above the National Average":

"A national survey by a school reform group has raised some serious questions about the significance of rising scores on standardized achievement tests. **The Friends for Education**, a West Virginia-based research group, says that the success often indicated by improving test scores is greatly exaggerated or, in some instances—practically meaningless."

—*The Washington Post*

"The public wants to feel safe from rigged scales or any other device to short-change the customer. It's time to adopt the same philosophy in education."

—Albert Shanker, *President, AFT*

"Is there a moral or ethical issue here? The fact that educators and test makers are making public statements that they know are misleading to parents and taxpayers would suggest that there is. Dr. Cannell, however, says that the issue goes beyond old-fashioned trust to one of justice."

—Edward B. Fiske, *The New York Times*

"A prediction: When grade-school students take achievement tests this spring, most of their scores will rise from a year ago. Now the bad news: Those scores will be virtually meaningless. So far, no one disputes John Jacob Cannell's remarkable conclusions, not even the testing companies themselves."

—Vincent Carroll, *The Rocky Mountain News*

"Lake Wobegon may never be the same. But at least the world of student testing will no longer be a world without logic."

—Mary Hatwood Futrell, *President, NEA*

FRIENDS FOR EDUCATION

600 Grand Blvd. NE, Albuquerque, NM 87106